Doctrine Formulation and Distrust

Toby J Heytens
DOCTRINE FORMULATION AND DISTRUST

Toby J. Heytens*

Legal scholars exhaustively debate the substantive wisdom of Supreme Court decisions and the appropriate methods for interpreting legal texts, but rarely consider the more pragmatic need to craft rules that will be faithfully implemented by the lower court judges who have the last word in the overwhelming majority of cases. Political scientists, in contrast, invest tremendous effort seeking to determine whether lower courts “comply” with Supreme Court directives, but find themselves unable to explain why their own studies generally find high levels of compliance. This Article argues that part of the answer lies in the Court’s ability to craft legal doctrines that both shape a trial court’s initial decision and increase the efficacy of appellate monitoring. After identifying numerous strategies for increasing lower court control, this Article argues that appreciating the links between them helps illuminate recent developments in three areas of public law – the constitutional law of punitive damages; the rules governing “officer suits” brought under 42 U.S.C. § 1983; and the concept of “reasonable” searches and seizures under the Fourth Amendment.

INTRODUCTION ............................................................................................................1
I. THE SUPREME COURT’S PROBLEM ........................................................................4
II. LEGAL DOCTRINE AS A TOOL FOR SUPREME COURT CONTROL ....................9
   A. Shape the Initial Decision .......................................................................................10
   B. Increase Appellate Oversight ..................................................................................19
III. THE STRATEGIES EMPLOYED ............................................................................25
   A. Modern Distrust of State Courts: The Constitutional Law of Punitive Damages .........26
   B. Reining in Federal Trial Courts: Constitutional Tort Suits .......................................34
   C. Switching the Directionality: “Reasonable” Searches and Seizures..........................44
CONCLUSION ..............................................................................................................50

* Associate Professor of Law, University of Virginia. This Article benefited greatly from presentations at the University of Virginia and Washington & Lee Law Schools. Thanks to Rick Bierschbach, Bruce Boyd, Paul Clement, Lynda Dodd, Amanda Frost, Brandon Garrett, Dave Hamilton, John Harrison, Rich Hynes, Jody Kraus, Clark Lombardi, Liz Magill, Michelle Morris, Greg Mitchell, Trevor Morrison, Dan Ortiz, George Rutherglen, Jim Ryan, Sarah Sawtelle, Steve Vladeck, Daniel Young, and David Zaring for helpful comments and suggestions, and to Kathleen Tarbox for excellent research assistance. This Article is dedicated to the memory of Judge Edward R. Becker. Having paraphrased Dean Ely’s title, I also borrow his dedication: “You don’t need many heroes if you choose wisely.” JOHN HART ELY, DEMOCRACY AND DISTRUST: A THEORY OF JUDICIAL REVIEW i (1980).
DOCTRINE FORMULATION AND DISTRUST

INTRODUCTION

Legal scholars exhaustively debate the substantive wisdom of Supreme Court decisions and appropriate methods for interpreting legal texts. Yet even scholars who emphasize the need to consider the more pragmatic process of translating first-order legal meaning into second-order legal doctrine have tended to neglect one critically important consideration: the need to craft rules that can and will be faithfully implemented by the lower court judges who have the last word in the overwhelming majority of litigated cases. “Judicial policies,” we lawyers too often forget, “do not implement themselves.”

In contrast, political scientists have invested great effort in determining whether lower courts have “complied” with Supreme Court directives, both during the Warren era and more recently. But committed as they generally are to the view that legal reasoning and legal doctrines are merely means by which judges seek to advance their policy aims, these scholars have been unable to advance a convincing explanation for why their own studies almost invariably find high levels of compliance.

This Article is different. Like the political scientists, it takes seriously the challenge the Supreme Court faces as it attempts to control lower court

---


2 Bradley C. Cannon & Charles A. Johnson, Judicial Policies: Implementation and Impact 1 (2d ed. 1999). Paul Gerwitz’s Remedies and Resistance, 92 Yale L.J. 585 (1983), is a notable exception. Because Professor Gerwitz dealt exclusively with desegregation decisions and examined resistance by both public and private actors, however, his focus was both narrower and much broader than mine.


4 See Kim, supra note 3, at 384.

5 See id. at 396-404 (reviewing political scientists’ explanations and identifying problems); see also Frank Cross, Appellate Court Adherence to Precedent, 2 J. Empirical Legal Stud. 369, 371-78 (2005) (same).

6 I recognize that “the Supreme Court” is a continuing collective body whose Members have no direct influence over the appointments process and little ability to control other Justices; I also recognize that the necessity of rendering group decisions imposes serious limitations on the Court’s ability to develop a coherent set of principles. See, e.g., Frank H.
behavior. But, like most legal scholarship, it also presumes that law actually matters. Building on those premises, this Article’s central arguments are that: (1) the Supreme Court has at its disposal a number of doctrinal tools that can be used to shape and direct lower court behavior; and (2) a number of recent developments suggest that attempts to control lower courts are alive and well.

The Supreme Court can seek to improve its control over lower court outcomes in two general ways. First, the Court can aim to channel a trial court’s initial decision. Second, it can increase monitoring by itself and other reviewing courts.

The Court has numerous techniques for influencing trial court decisions. Because complicated or open-ended standards increase the risk of good-faith misunderstandings and create opportunities for disguising deliberate non-compliance, the Court may be better served by laying down simple rules whose application depends on only a few factors. The Court can also bar trial courts from relying on criteria it deems irrelevant or too difficult to verify on appeal—such as, for example, the presence or absence of a particular state of mind—and it can enforce these rules of exclusion by imposing upon trial courts duties.

Easterbrook, *Ways of Criticizing the Court*, 95 HARV. L. REV. 802 (1982); Lewis A. Kornhauser & Lawrence G. Sager, *Unpacking the Court*, 96 YALE L.J. 82 (1986). Despite the inevitable risks of distortion, I will, for ease of clarity, largely ignore group dynamics within the Court itself, citing as only partial defense the fact that scholars who have examined the nature of the Court as a group decisionmaker have generally assumed away the existence of lower courts entirely. *See*, e.g., Lewis A. Kornhauser, *Modeling Collegial Courts I: Path Dependence*, 12 INT’L REV. OF LAW & ECON. 169, 170 (1992) (assuming “the judicial system consists of a single court that decides all cases”).


8 Consideration of whether it is normatively justifiable for the Justices to seek to control lower court behavior, however, must await another day. *See infra* text accompanying notes 298-300.

9 As is generally the case with taxonomies, these two categories are not entirely distinct. *See infra* note 41 and accompanying text.

10 Both of these strategies are, at least in principle, available to any appellate tribunal seeking to increase its control over trial court outcomes. I focus on the Supreme Court for two reasons. First, the literatures I seek to engage have tended to focus on that Court. Second, the situation confronting intermediate appellate courts is complicated by the fact that their efforts to secure greater control over trial court outcomes are themselves subject to override by the Supreme Court. *See*, e.g., KSR Int’l Co. v. Teleflex, Inc., 127 S. Ct. 1727, 1734-35, 1739-43 (2007) (rejecting attempt by Federal Circuit to specify in greater detail manner in which federal district courts should decide whether invention was “obvious” for purposes of federal patent laws).
to explain their decisions. Finally, the Court can nudge trial courts towards a favored result or away from a disfavored one by adjusting the background legal baseline (such as whether a particular type of law is presumed unconstitutional) or by adjusting the level of proof necessary to dislodge it.

On the monitoring side, there is likewise a great deal the Court can do. It can expand the availability of appellate review by creating exceptions to the final judgment rule – a review-limiting doctrine that generally bars an appeal until a trial court has issued a decision that “ends the litigation on the merits and leaves nothing more for the court to do but execute the judgment.” 11 It can increase the intensity of appellate monitoring by characterizing certain issues as presenting questions of law rather than questions of fact and by otherwise adopting non-deferential standards of appellate review. And the Supreme Court can enforce the restrictions it imposes on trial courts’ initial determinations by adopting appellate presumptions – such as one that all trial court errors have an adverse effect on a verdict – to be applied when those directives are not followed.

All of the strategies just mentioned share certain overlapping downsides. Some techniques – including, to varying degrees, all of them that seek to control a trial court’s initial decision – can be difficult to use well without possessing the sort of information that the Justices may have a difficult time gathering. Other strategies – most notably rules and significant adjustments to the required proof level – risk generating new undesirable outcomes even as they work to prevent certain others. Use of some strategies raises concerns about legitimacy, and still others can be difficult to implement or costly to apply and enforce.

In addition, no strategy for increasing Supreme Court control over lower court outcomes is foolproof. All of the techniques discussed in this Article rely on the assumption that basic norms regarding appropriate judicial behavior (including, most especially, the one that lower court judges will not simply lie about the bases for their decisions) 12 have at least some constraining effect. 13 In addition, sheer numbers alone will always preclude fully effective monitoring of trial courts by intermediate appellate courts or of those appellate courts by the Supreme Court itself. 14 Far from making the extravagant claim

12 See generally David L. Shapiro, In Defense of Judicial Candor, 100 HARV. L. REV. 731, 750 (1987) (arguing that “the fidelity of judges to law can be fairly measured only if they believe what they say in their opinions and orders”).
14 See infra notes 32-35 and accompanying text.
that legal doctrines alone can generate perfect Supreme Court control of lower court outcomes, this Article will simply advance the more modest one that they represents one underappreciated tool for doing so.

The remainder of this Article is organized as follows. Part I will describe the challenges the Supreme Court faces in attempting to control lower court outcomes. Part II will turn to legal doctrine, describing a number of ways in which the Court can use it to shape and direct lower court behavior and discussing the relative advantages and disadvantages of each. Part III will apply the insights of Part II to recent developments in three areas of public law – the constitutional law of punitive damages, the rules governing “officer suits” against state officials, and the concept of “reasonable” searches and seizures under the Fourth Amendment. These examples, it will argue, suggest that Supreme Court attempts to control lower court behavior are very much alive, and illustrate why the Court might sometimes choose one strategy for doing so over another. A brief Conclusion will identify some implications of this analysis and suggest areas for future research.

I. THE SUPREME COURT’S PROBLEM

“If men were angels,” James Madison wrote, “no government would be necessary.”15 Something similar is true of the relationship between the Supreme Court and lower courts. If lower court judges were nothing more than “the simple (and perhaps simple-minded) enforcers of the Supreme Court’s dictates”16 – or if appellate review was perfectly accurate, costless, and instantaneous – there would be no need for the Justices to consider the possibility of good faith error or deliberate lower court resistance when formulating legal doctrines. But that is not the world in which we live. Indeed, as this Part explains, the Supreme Court cannot simply assume that lower courts will implement the Court’s decisions in a manner that a majority of Justices would view as correct.17

Although the degree to which lower courts do or do not follow Supreme Court precedent is “[u]ltimately . . . an empirical” question,18 any fully successful demonstration of the risk of lower court non-compliance must, at

\[\text{15 THE FEDERALIST No. 51 (James Madison).}\]


\[\text{17 Because this Article views matters from a top-down, Supreme Court-focused perspective, its sole test for assessing the correctness of a lower court decision is whether it is the outcome that would be reached by a majority of fully informed Justices. This Article offers no method for how the Justices themselves should decide cases, nor any metric for assessing whether their decisions are correct.}\]

\[\text{18 Cross, supra note 5, at 398.}\]
least for now, be largely conceptual rather than entirely empirical. Fortunately, such a framework already exists, one originally developed to analyze economic organizations but since applied to a number of public actors, including judges: principal-agent theory.

Broadly speaking, a principal-agent relationship exists whenever one person or group (a principal) must rely on others (agents) to accomplish the principal’s goals. Reliance on others has two disadvantages vis-à-vis a do-it-yourself approach. First, the agent may be less skilled at performing her assigned tasks than the principal would have been, be it because of less training or ability, workload, or lack of sufficient direction from the principal. Second, principals and agents will likely inevitably have at least somewhat different preference structures. These twin disadvantages, in turn, create “agency costs” — the deadweight loss created by the fact that agents are inevitably somewhat imperfect proxies for their principals.

---

19 Because the Supreme Court has almost total freedom in deciding what cases it will hear, see Richard H. Fallon, Jr., et al., Hart and Wechsler’s The Federal Courts and the Federal System 468, 1580 (5th ed. 2003) [hereinafter Hart & Wechsler], one cannot take its raw reversal rate as representing the Justices’ views about overall lower court behavior. And despite a number of high profile counter-examples involving the white-hot issues of Federal versus State supremacy, see, e.g., Martin v. Hunter’s Lessee, 14 U.S. 304 (1816); Tarble’s Case, 80 U.S. 397 (1871); race, see, e.g., J.W. Peltason, Fifty-Eight Lonely Men 7 (1961); 22 Negroes Fined in Bus Bias Case, N.Y. Times, Mar. 22, 1957, at 14 (discussing Alabama state court judge who refused to follow Supreme Court ruling that segregation of municipal busses was unconstitutional); and religion, see, e.g., Jaffree v. Bd. of Sch. Comm’rs of Mobile County, 554 F. Supp. 1104 (S.D. Ala.) (holding, in conceded contravention of decades of Supreme Court precedent, that “the establishment clause of the first amendment to the United States Constitution does not prohibit [a] state [government] from establishing a religion”), instances in which lower court judges claim to be doing anything other than making a good faith attempt to apply existing high court authority are basically nonexistent. Accordingly, one could not obtain a fully accurate empirical measure of the extent of lower court non-compliance without devising some method for determining what a majority of Justices would think about cases the Court never hears and then applying that technique to some appropriately selected subset of all lower court rulings. See Barry Friedman, Taking Law Seriously, 4 Persp. on Pol. 261, 271 (2006) (highlighting dangers with ignoring unpublished opinions when attempting to draw generalizable conclusions about judicial behavior).


21 Of course, one reason principals retain agents is because they possess skills or knowledge the principal lacks, though it may also be that the scope or complexity of the activity in which the principal seeks to engage is beyond the capacity of any single person. See, e.g., Terry M. Moe, The New Economics of Organization, 28 Am. J. Pol. Sci. 739, 756 (1984).

22 See, e.g., id. at 756.
If one conceives of the Supreme Court as a principal and lower courts as its agents, it is clear that the Court lacks the ability to employ many of the methods traditionally used to address agency costs. Take agent selection. Although Justices have sporadically advised Presidents about the appointment of some federal judges, they have no direct role in such appointments and no right to block the appointment or promotion of any particular lower court judge. Instead, state and lower federal court judges are appointed by a host of independent political processes that may well select for things other than high judicial competence, ideological compatibility with Supreme Court preferences, or willingness to subordinate one’s personal views to those of one’s judicial superiors. What is more, because appointments tend to be made on a rolling basis, even a Justice who is confident that most lower court judges appointed at the same time as her share her general views may still worry that those whose service pre- or post-dates hers may not merit that description. And finally, because lower court judges generally hear cases either alone or in panels of less than the size of the full court – and are assigned to do so via processes over which Supreme Court Justices have no control – even a relatively small number of unreliable lower court agents will have the potential to affect the outcome of a great many cases.

Not only do the Justices lack the ability to select their lower court agents, their ability to incentivize what they would regard as good performance is limited at best. Judges on a given lower court are generally paid the same amount, and their level of compensation is determined by the jurisdiction

23 Although this characterization is fairly standard among political scientists, see Kim, supra note 3, at 386 n.13 (citing sources), it is nonetheless hardly free of controversy. In particular, Professor Pauline Kim has argued that viewing the relationship between the Supreme Court and lower courts through a principal-agent framework often rests on a series of unarticulated and difficult-to-defend normative presuppositions regarding the nature of the judicial hierarchy. See id. at 434-441. My use of the principal-agent framework, however, is for the limited purpose of explaining why a Justice who wants her own preferences followed has cause to worry about lower court non-compliance. Of course, whether it is legitimate for a Justice to act on such concerns is a different matter entirely. See infra text accompanying notes 295-96.


25 See U.S. CONST. art. II, § 2, ¶ 2, cl. 2.

26 See generally Nina Totenburg, Will Judges Be Chosen Rationally?, 60 JUDICATURE 92 (1976) (describing selection process for federal judges as it existed in middle part of the twentieth century).

27 See 28 U.S.C. § 46(a) (“Circuit judges shall sit on the court and its panels in such order and at such time as the court directs”); § 137, ¶ 1 (authorizing Federal District Courts “having more than one judge” to divide cases “as provided by the rules and orders of the court”).
(federal or state) in which they sit and their position within that judicial hierarchy rather than their level of compliance with Supreme Court decisions. The Justices cannot fire, demote, transfer, or financially penalize bad lower court judges any more than they can promote or give raises to good ones. They cannot order that a specific lower court judge be precluded from hearing certain kinds of cases, even if the judge in question has repeatedly shown herself to be an unreliable executor of Supreme Court policy. In theory, the Justices could hold a particularly recalcitrant lower court judge in contempt, but the Court has never done so and it has threatened it only a handful of times. Reversals accompanied by sharp criticism may have some effect on lower court behavior, as may affirmances accompanied by effusive praise, but the success of these techniques depends on the sensitivity of particular lower court judges to Supreme Court feedback (and the public attention that often accompanies it), as well as the likelihood of monitoring in any given case.

Perhaps the biggest problem, however, is that the Court simply lacks the capacity to monitor more than a tiny fraction of lower court decisions. First, there is the matter of raw numbers. The Supreme Court itself has never issued more than 298 merits decisions in a single year, and in recent years it has generally issued around 80. In 2004 alone, nearly 1,000 federal district judges

---

28 See 28 U.S.C. § 44(d) (setting a single salary for federal circuit judges); § 135 (federal district judges); see also Richard A. Posner, What Do Judges Maximize? (The Same Thing Everybody Else Does), 3 SUP. CT. ECON. REV. 1, 2 (1993) (“[A]lmost the whole thrust of the rules governing compensation and other terms and conditions of judicial employment is to divorce judicial action from incentives.”).

29 I mean “bad” solely in the sense of being poor agents of the Supreme Court.

30 The most recent example was in 1969, when the Supreme Court considered a motion for an order to show cause why an Alabama state judge should not be held in contempt for violating an order issued by the Court in a voting rights case. Although Justices Douglas and Harlan would have granted the motion, a majority voted to stay consideration pending a determination of whether the same judge had violated an earlier order by federal district court. See In re Herndon, 394 U.S. 399 (1969).

31 See WALTER MURPHY, ELEMENTS OF JUDICIAL STRATEGY 104-05 (1964). At least one study, however, has found that a district judge’s “reversal rate” has no effect on likelihood of promotion. See Richard S. Higgins & Paul R. Rubin, Judicial Discretion, 9 J. LEGAL STUD. 129, 137 (1980).

32 The year was 1886. See LEE EPSTEIN ET AL., SUPREME COURT COMPENDIUM tbl 3-2 (3d ed. 2002).

disposed of more than 300,000 cases and nearly 30,000 state trial court judges resolved nearly 19,000,000.\textsuperscript{34} It is true that the Justices can get help from the thirteen United States Courts of Appeal and various intermediate state appellate courts, which handed down 56,381 and 240,827 decisions in 2004, respectively.\textsuperscript{35} But even these other appellate courts, the numbers make clear, cannot possibly review anywhere near all trial court decisions, and, at any rate, Supreme Court reliance on other appellate courts to monitor trial courts simply introduces another principal-agent relationship – one between the Supreme Court and the members of those other appellate courts.

To make matters worse, factors other than sheer numbers further hinder appellate courts’ ability to monitor trial courts and the Supreme Court’s ability to monitor either. Appellate courts cannot review any lower court ruling unless the losing party files an appeal (or, in the case of the Supreme Court and many state courts of last resort, multiple appeals). Even if a litigant who receives an unfavorable ruling from a lower court determines that an appeal is worth the cost, a number of generally applicable doctrines – including, most notably, the final judgment rule and deferential standards of review – operate to limit the effectiveness of appellate monitoring.\textsuperscript{36} And even if non-compliance is detected, there is little tangible an appellate court can do other than change the outcome of that particular case, and often the necessity of further fact-finding or the availability of other bases for decision will conspire to prevent even that.\textsuperscript{37}

---

\textsuperscript{34} See 2005 Judicial Facts and Figures, tbl. 1.1 (664 active and 291 senior federal district court judges); id. tbl. 6.1 (federal district courts terminated 317,382 cases), available at http://www.uscourts.gov/judicialfactsfigures/; Shauna M. Strickland, \textit{State Court Caseload Statistics}, 2003, fig. G (11,349 state trial court judges sitting on general jurisdiction courts and 18,161 sitting on limited jurisdiction courts); tbl 1, summary section for all trial courts (9,833,278 civil and 9,240,776 criminal cases disposed of by state trial courts), available at http://www.ncsconline.org/d_research/csp/2005_files/State\%20Court\%20Caseload\%20Statistics\%202005.pdf. The state court numbers are misleadingly high, however, because the only state court cases the Supreme Court may review are those that include a contested issue of federal law. See \textit{Murdock v. City of Memphis}, 87 U.S. 590 (1875).

\textsuperscript{35} For the United States Courts of Appeal, see 2005 Judicial Facts and Figures, supra note 34, at tbl. 2.1. For state courts of last resort, see Strickland, supra note 34, at tbl. 10, summary section for all appellate courts.

\textsuperscript{36} See infra notes 94-100 and accompanying text.

\textsuperscript{37} See \textit{Hart \& Wechsler}, supra note 19, at 517. For dated, but still fascinating examinations of the fate of Supreme Court decisions on remand to state courts, see Note, Final Disposition of State Court Decisions Reversed and Remanded by the Supreme Court, October Term, 1931, to October Term 1941, 55 \textit{Harv. L. Rev.} 1357 (1942), Note, Evasion of Supreme Court
II. LEGAL DOCTRINE AS A TOOL FOR SUPREME COURT CONTROL

So what’s a poor Supreme Court Justice who wants to exert at least some control over lower court outcomes to do? She might daydream about the power to hire and fire lower court judges, or even to set their salaries. She might wish for the ability to reorganize certain structural features of the United States judicial system, perhaps eliminating the state courts or changing the number of layers of appellate review or the allocation of resources between trial and appellate courts. But let’s take current institutional arrangements as given: There is actually quite a lot the Justices can do to shape and direct lower court behavior.

The source of this power lies in the Court’s ability to formulate and later adjust legal doctrines based on the perceived risk of lower court noncompliance. Speaking broadly, a Justice who wants to limit the number of outcomes with which she disagrees can employ one of two approaches: (1) aim to shape and direct the underlying merits determination; or (2) lower barriers to effective monitoring by the Court itself or trusted appellate courts.

There is obviously a great deal of overlap between these approaches. Techniques that channel the primary merits decision can also facilitate monitoring by making non-conforming decisions easier to spot. And strategies that focus on removing obstacles to effective monitoring can reduce the number of initial errors if, as commonly believed, lower court judges generally dislike being reversed. In fact, one of this Part’s key objectives is to demonstrate how the various strategies that the Supreme Court can use to increase its control over lower court outcomes overlap with, and are, at least to a certain extent, substitutes for, one another.

Mandates in Cases Remanded to State Courts Since 1941, 67 HARV. L. REV. 1251 (1954), and Note, State Court Evasion of United States Supreme Court Mandates, 56 YALE L.J. 574 (1947).

38 See, e.g., Burke Shartel, Federal Judges—Appointment, Supervision, and Removal—Some Possibilities Under the Constitution, Part I, 28 MICH. L. REV. 485, 488 (1930) (arguing it would be constitutional to have “[a]ppointment of inferior federal judges by the judiciary branch itself”); Burke Shartel, Federal Judges—Appointment, Supervision, and Removal—Some Possibilities Under the Constitution, Part III, 28 MICH. L. REV. 870, 875 (1930) (asserting it would be constitutional to authorize federal judiciary to remove “unfit” judges).


41 See, e.g., Calvin Magruder, The Trials and Tribulations of an Intermediate Appellate Court, 44 CORN. L.Q. 1, 7 (1958) (“I don’t enjoy getting reversed more than any other judge . . . ”).
A. Shape the Initial Decision

A trial court can reach a result other than the one a majority of Justices would desire for one of two main reasons. First, the trial court may misunderstand or misapply the standards laid down by the Supreme Court and thus commit legal error. Second, even perfectly faithful application of the Supreme Court’s directives may sometimes produce what a majority of Justices would regard as a bad outcome. (To cite a classic non-judicial example of what most people would regard as a bad outcome: A prohibition on lying would, if strictly followed, require telling a killer the location of his intended victim.) To make matters even more complicated, the Justices may see certain kinds of bad outcomes as worse than others, as reflected by the adage that it is “[b]etter that ten guilty persons escape, than that one innocent suffer.” Accordingly, perhaps the biggest challenge facing a Justice who seeks to influence lower court behavior is how to craft doctrines that will minimize the number of legal errors while simultaneously avoiding at least the most objectionable sorts of bad outcomes.

The risk an initial decisionmaker will commit legal error is largely a function of what Professor Maurice Rosenberg termed “primary” discretion – the extent to which the trial court is empowered to make “a wide range of choices . . . free from the constraints which characteristically attach whenever legal rules enter the picture.” The most obvious way to reduce the number of legal errors, therefore, is to limit this discretion by placing restrictions on the number of factors a trial court may consider or the range of outcomes it may reach.

The most familiar strategy for reducing an initial decisionmaker’s primary discretion involves use of rules rather than standards. Although the distinction invariably grows fuzzy around the margins, standards typically call for a

---

42 In reality, the point made in the text applies not only to trial courts, but also intermediate appellate courts engaged in non-deferential review of earlier trial court decisions. Because the next Section focuses on doctrines directed exclusively at reviewing courts, I will for clarity’s sake generally speak in this Section of tactics that aim to control a trial court’s initial decision.

43 Again, by “bad outcomes,” I mean simply those a majority of Justices would regard as undesirable.

44 See, e.g., Immanuel Kant, On a Supposed Right to Lie Because of Philanthropic Concerns, in GROUNDING FOR THE METAPHYSICS OF MORALS WITH ON A SUPPOSED RIGHT TO LIE BECAUSE OF PHILANTHROPIC CONCERNS 63-64 (James W. Ellington trans., 1993) (1785) (discussing argument made by Benjamin Constant).


46 See Maurice Rosenberg, Judicial Discretion of the Trial Court, Viewed From Above, 22 SYR. L. REV. 635, 637 (1971).
situation-specific, all-things-considered type inquiry whereas rules direct consideration of a relatively small number of factors and prescribe rigid consequences based on their presence or absence. To use a familiar example: “No one shall drive faster than is safe under the circumstances” is a standard; “Speed Limit 55” is a rule.

The existing literature about the choice between rules and standards has generally neglected its implications for the relationship between the Supreme Court and lower courts. That said, it is easy to see why a Justice concerned about lower court compliance may often prefer to set forth legal doctrines in rule form. As a number of prominent scholars have pointed out, perhaps the fundamental characteristic of decisionmaking by rule is its ex ante character: the rule-maker performs a before-the-fact assessment of all relevant considerations and then declares what consequences should follow in a wide variety of circumstances. Because the Supreme Court is so rarely the initial decisionmaker in any case, its own power is enhanced – and lower courts’ primary discretion is reduced – when the Justices formulate legal doctrines in such a way as to make as many of the hard choices as possible at the doctrine-creation phase. In addition, reducing the number of facts that may be considered and restricting an initial decisionmaker’s options (both

---


49 See Hart & Wechsler, supra note 19, at 268-318 (discussing the Court’s “original” jurisdiction).

50 See Schauer, supra note 47, at 159 (noting that “[r]ules operate as tools for the allocation of power”). For the same reason, rules may seem attractive to a Justice who wants to influence the outcome of future Supreme Court decisions, see, e.g., Antonin Scalia, The Rule of Law as a Law of Rules, 56 U. CHI. L. REV. 1175, 1179-80 (1989), though there one must confront the impressive body of social science research suggesting that Supreme Court Justices are, at very least, highly resistant to efforts at doctrinal control. See generally Jeffrey A. Siegel & Harold J. Spaeth, The Supreme Court and the Attitudinal Model Revisited (2002).
characteristics of rules) will tend to reduce the number of good-faith mistakes and facilitate more effective monitoring by reviewing courts.\textsuperscript{51}

So why not use rules all of the time? The answer is because rules – like every other strategy for increasing Supreme Court control over lower court outcomes – have costs as well as benefits. Indeed, in the case of rules, the downsides are basically the advantages in reverse, and they will often be substantial.

For one thing, Justices can confront serious \textit{informational barriers} when attempting to formulate rules. Because rules seek, in effect, to resolve certain categories of cases before they arise, it can hard to create good ones without extensive information.\textsuperscript{52} A Justice may lack adequate understanding of the current state of affairs in the lower courts, and find it hard to predict the precise consequences of various possible rules on lower court behavior. Even if these problems can be overcome, the Justice may find it difficult to express her own preferences in terms of just a few basic criteria, to say nothing of the challenges posed by trying to anticipate what she will think about situations that have not yet even arisen.\textsuperscript{53}

A related drawback of rules – one they share with all strategies that rely on contracting a trial court’s primary discretion – is that they tend to reduce the number of legal errors only at the cost of generating \textit{new bad outcomes}. Rules gain their power to control lower court decisionmaking by constricting the range of relevant information and limiting a trial court’s ability to take account of situation-specific equities in choosing what course to follow. By doing so, however, rules almost inevitably generate situations in which the outcome dictated by the rule’s application does not further – and may affirmatively

\textsuperscript{51} See, e.g., Ashutosh Bhagwat, \textit{Separate But Equal?: The Supreme Court, the Lower Federal Courts, and the Nature of the “Judicial Power”}, 80 B.U. L. REV. 967, 990 (2000) (“Because lower court judges are rarely willing to explicitly flout binding precedent, . . . if the Supreme Court were to establish and explicate clear, doctrinal rules with determinate consequences, such rules might be quite effective in actually binding lower court decision-making and limiting the avoidance of precedent.”); Russell B. Korobkin, \textit{Behavioral Analysis and Legal Form: Rules vs. Standards Revised}, 79 OR. L. REV. 23 (2000) (noting that even if initial adjudicators can engage of bad-faith manipulation of rules as easily as they can with standards, “errors in rule application will tend to be more obvious and, therefore, more susceptible to correction on appeal”).

\textsuperscript{52} See, e.g., Ehrlich & Posner, \textit{supra} note 47, at 267 (“Obtaining and correctly evaluating information concerning the various combination of events or circumstances under which the general standard that the set of rules is designed to implement should be activated are costly.”).

\textsuperscript{53} See H.L.A. Hart, \textit{The Concept of Law} 125 (1961) (identifying “relative indeterminacy of aim” as one of “two connected handicaps” that afflicts efforts “to regulate unambiguously and in advance, some sphere of conduct by means of general standards to be used without further official direction on particular occasions”).
contradict— the purposes that led to the rule’s creation in the first place, and
the frequency of such occurrences will generally increase as the rule grows
simpler and more determinate.54 Assuming that at least one reason why a
Justice would seek to control lower court behavior is to maximize the overall
number of what she would regard as correct outcomes, the prospect
generating new bad ones would seem to be cause for serious concern.55

A third downside is largely unique to the use of rules. Rules often have at
least two audiences—an out-of-court actor whose conduct the rule governs
and the officials who will later decide whether the rule has been violated.56
Because there is rarely perfect “acoustic separation” between these two
audiences,57 there may be situations in which even a rule that is well-designed
to control lower court decisionmaking may have undesirable impacts on non-judicial
actors. Take, for example, the Supreme Court’s decision in County of Riverside v.
McLaughlin,58 which addressed how best to operationalize the principle that a
person detained without a warrant must receive a judicial hearing “promptly
after arrest.”59 Although the Court’s rule-like holding that hearings held more
than 48 hours after arrest are presumptively invalid undoubtedly limits a trial
court’s primary discretion, it may also, the Court’s contrary protests
notwithstanding,60 be seen as giving executive branch officials free rein to
decide when within the first 48 hours such hearings should be held.

Finally, use of rules may sometimes raise concerns about legitimacy. Justice
Scalia may be correct that it is often possible to construe “even the most vague
and general text” as having “some precise, principled content.”61 In some
situations, however, the controlling legal text may be phrased so broadly as to
make it difficult for the Court to characterize creation of a clear rule as an act
of judicial “interpretation” and doing so may instead, at least to some people,
begin to resemble an act of impermissible judicial “legislation.”62

54 See Ehrlich & Posner, supra note 47, at 268; Sunstein, supra note 48, at 990.
55 A Justice might find rules attractive for reasons other than maximizing the number of
correct outcomes, such as ensuring similar outcomes for similarly situated litigants.
56 See Ehrlich & Posner, supra note 47, at 261; Sunstein, supra note 48, at 960.
57 See Meir Dan-Cohen, Decision Rules and Conduct Rules: On Acoustic Separation in Criminal
59 See id. at 52 (quoting Gerstein v. Pugh, 420 U.S. 103, 125 (1975)).
60 See id. at 56.
61 Scalia, supra note 50, at 1183.
62 The classic debate here is the one regarding the legitimacy of Miranda v. Arizona, 384
Question of Article III Legitimacy, 80 NW. U. L. REV. 100 (1985) (Miranda illegitimate), with David
and Berman, supra note 1, at 114-165 (same). Although my own sympathies are with
For all of these reasons, rules may not always be an attractive option for a Justice seeking to limit trial courts’ primary discretion. What should not be overlooked, however, is that there are a number of other possible options.

A second strategy involves what I will call rules of exclusion. As explained earlier, rules operate by taking a relatively small number of facts and making their presence or absence outcome determinative. Rules of exclusion are more limited: they tell an initial decisionmaker not to consider certain facts. To cite a fairly well-known example: When the Supreme Court held in *North Carolina v. Pearce* that although trial courts may impose a longer sentence following a second trial for any number of reasons, they may not do so to punish the defendant for having taken a successful appeal, it was announcing a rule of exclusion. Another example would be situations in which the Supreme Court has barred consideration of a party’s subjective intent.

As a strategy for increasing lower court control, rules of exclusions’ chief selling point is that they permit the Supreme Court to obtain at least some of the benefits of rules while minimizing the downsides. Rules of exclusion can increase the number of initially correct lower court decisions by barring resort to considerations that the Justices have concluded will, more often than not, lead lower courts in the wrong direction, or are too likely to lead to a particular type of disfavored result. Rules of exclusion can also facilitate monitoring by precluding trial courts from relying on variables whose presence or absence is particularly hard to verify on appeal or whose precise significance would be hard to confirm or refute based on a cold record. Although they do require some effort to formulate and may generate some new bad outcomes if the forbidden consideration has even a slight connection to the underlying purpose of the inquiry, rules of exclusions’ more limited nature means that they will generally require less information to create than rules and the negative

---

63 Professor Ehlich and Judge Posner would say what I call rules of exclusion are simply comparatively less precise rules. See Ehlich & Posner, supra note 47, at 258 (“A rule withdraws from the decision maker’s consideration one or more of the circumstances that would be relevant to decision according to a standard.”). As explained in the text, however, there are several important distinctions between rules and rules of exclusion when used as tools of ensuring greater Supreme Court control over lower court decisionmaking.


65 See id. at 723-25.

66 See infra notes 207-10 & 271-72 and accompanying text.

67 See Schauer, supra note 47, at 150 (“Often we fear that some class of decision-makers, whether through unconscious ill-will, cannot be trusted to take certain types of factors into account.”).
impact of excluding a single variable from consideration will generally be less severe than making the entire decision turn on just one or two facts. Finally, rules of exclusion will rarely suffer from serious legitimacy problems, because the Court can simply construe the underlying constitutional or statutory provision as making the excluded fact legally irrelevant.68

Rules of exclusion, of course, are no panacea. For one thing, the very characteristic that will often make them less problematic than rules – that is, their more limited nature – will also make them relatively less effective at securing lower court decisions with which a majority of Justices would agree.

Rules of exclusion also have another downside vis-à-vis rules, one they share with virtually every other technique for enhancing Supreme Court control over lower court decisionmaking. As compared with rules, rules of exclusion can be both hard to implement and costly to apply and enforce. It is comparatively easy to determine whether the evidence in the record was sufficient to support a trial court’s conclusion that a driver was going 90 miles per hour in a 55 zone and whether the fine it imposed was the one set by the governing statute. But absent an on-the-record declaration, it will often be both challenging and time-consuming to deduce whether a trial court relied on a forbidden consideration in reaching an otherwise permissible result.

One way to address this last problem is to couple rules of exclusion with a third strategy and to impose a duty of explanation. Indeed, the Supreme Court did precisely that in Pearce, the case that held that trial courts may not seek to punish defendants for taking successful appeals.69 Citing the need “to assure the absence of such a motivation,” Justice Stewart’s majority opinion stated that a trial judge who decides to impose an enhanced sentence must make an on-the-record declaration of her reasons.70

The main advantages of this strategy should be obvious. Duties of explanation require little information to adopt. And, at least standing alone, they are extremely unlikely to generate new bad outcomes.

But duties of explanation too have their downsides. For one thing, the explanation may well not be forthcoming unless the duty is backed up by an appellate presumption, a review-expanding technique discussed in the next Section,71 and one possessing its own problems. In addition, it will often be easy for a trial judge who seeks to avoid reversal to say that she is not relying

68 See, e.g., infra notes 207-10 and accompanying text (discussing decisions barring lower courts from considering a defendant’s subjective intent in ruling on motions for summary judgment based on qualified immunity).
69 See supra notes 64-5 and accompanying text.
71 See infra notes 108-15 and accompanying text.
upon any forbidden considerations, and hard for a reviewing court to tell if she is lying. And if all that is not enough, many people view judicially imposed duties of explanation as having serious legitimacy problems.\textsuperscript{72} Although seemingly everyone agrees that the Supreme Court may prescribe tests for determining whether a given governmental act violates the Constitution, a number of Justices have argued that the Court simply has no legitimate authority to tell state courts how they must go about explaining their decisions.\textsuperscript{73}

What if the Supreme Court wants to influence the course of initial trial court decisionmaking, but concludes that the techniques discussed so far would be impractical, unwise, or ineffective if imposed alone? At least two options remain.

A fourth strategy for controlling a trial court’s primary discretion is to adjust the baseline presumption. By “baseline presumption,” I mean the background principle that controls in the absence of legally cognizable evidence to the contrary or when the evidence is in equipoise.

Baseline presumptions can be broad and general, or narrow and context-specific. To cite an example of the former: the baseline presumption in all criminal litigation and at the threshold of virtually all civil suits between private parties is that the defendant is not liable. In contrast, the baseline presumption can ping-pong back and forth several times: governmental activity is generally presumed constitutional,\textsuperscript{74} unless it implicates constitutionally protected rights,\textsuperscript{75} unless the person whose rights are being affected is in prison,\textsuperscript{76} unless the right in question is the one to be free from purposeful race discrimination.\textsuperscript{77}

\textsuperscript{72} A number of statutes and Federal Rules impose duties of explanation on federal trial judges. See, e.g., 18 U.S.C. § 3553(c) (directing federal district court charged with sentencing criminal defendant “in open court the reasons for its imposition of the particular sentence”); FED. R. CRIM. P. 12(d) & (f) (directing federal district courts to make express findings of fact when ruling on pre-trial suppression motions). In addition, the Fourth Amendment imposes a duty of explanation on federal and state judges alike. See U.S. CONST. amend. IV (stating all “warrants” must “particularly describ[e] the place to be searched, and the person or things to be seized.”)


\textsuperscript{74} See Fairbank v. United States, 181 U.S. 283, 285 (1901).
\textsuperscript{77} See Johnson v. California, 543 U.S. 499 (2005).
Standing alone baseline presumptions have both limited downsides and a limited capacity to influence lower court decisionmaking. Unlike rules or rules of exclusion, setting a baseline presumption does not require the Court to be able to specify in advance that any particular facts may – or may not – be considered. In addition, because a legal system simply cannot function without ways for deciding what happens absent proof to the contrary, 78 neither the difficulty of choosing the appropriate presumption nor the risk that picking the wrong ones will generate additional bad outcomes can justify – or even make possible – failure to select one. For similar reasons, it is difficult to lodge legitimacy objections against use of presumptions (at least of the “rebuttable” kind), 79 though there may be situations in which either a controlling legal text or widespread and longstanding practice would make it difficult to justify setting the presumption in one particular way. And finally, although there can be situations in which it can be difficult to figure out which of several possible presumptions should apply, 80 once the appropriate presumption is identified, it is comparatively easy for trial courts to assess, and appellate courts to monitor, whether there is no evidence to dislodge it.

At the same time, however, baseline presumptions alone are a fairly limited tool for securing greater Supreme Court control over trial court decisionmaking. It is a comparatively rare situation in which there is literally no evidence that is even arguably relevant with respect to an important point in litigation. And it may be even rarer to find a situation in which the evidence that does exist is self-evidently in equipoise.

Baseline presumptions can become quite a bit more effective, however, when paired with a fifth and final strategy for shrinking a trial court’s primary discretion: raising the level of proof necessary to justify a departure from the presumptive state of affairs. Sometimes these adjustments are overtly framed in terms of the burden of proof, such as the requirement of clear and

78 Cf. Berman, supra note 1, at 10 (explaining that because court’s invariably “lack[] unmediated access to the true fact of the matter,” they must by necessity adopt certain “decision rules” to help them deal with epistemic uncertainty).

79 One of Miranda’s leading academic critics argued that the Supreme Court lacked authority to establish “conclusive” presumptions, though he viewed “rebuttable” presumptions as constitutionally unproblematic. See Grano, supra note 62, at 141 & n.271, 142-48, 156 (1986) (rebuttable presumptions legitimate); Joseph D. Grano, Miranda’s Constitutional Difficulties: A Reply to Professor Schulhofer, 55 U. Chi. L. Rev. 174, 179 (1988) (conclusive presumptions illegitimate). For criticisms of this distinction, see Berman, supra note 1, at 136-54, and Strauss, supra note 62, at 191-95.

80 Compare Johnson v. California, 543 U.S. 499, 505-16 (2005) (O’Connor, J., for the Court), with id. at 528-532 (Thomas, J., dissenting) (disagreeing whether a case involving racial segregation in prisons was governed by the presumption that race-based classifications are unconstitutional or the presumption that prison regulations are constitutional).
convincing evidence in certain libel cases\textsuperscript{81} or proof beyond a reasonable doubt in criminal trials.\textsuperscript{82} Other times, they are framed in terms of levels of scrutiny, such as the requirement that a party who challenges the constitutionality of a statute regulating economic activity must demonstrate that there is no set of facts that could even arguably justify it.\textsuperscript{83}

As a strategy for limiting a trial court’s primary discretion, the chief upsides of adjusting the level of proof should be largely familiar by now. Raising proof requirements should push all but the most recalcitrant trial judges in the favored direction and ease monitoring by reviewing courts.\textsuperscript{84} Not only can this technique thus be particularly useful to a court seeking to avoid certain kinds of particularly disfavored outcomes – such as, for example, convicting the innocent – but there are even situations in which increasing the required level of proof “can, perhaps counterintuitively, reduce” the total number of undesirable outcomes.\textsuperscript{85} Finally, unlike rules, adjusting the level of proof does not require specifying all relevant considerations in advance, and, unlike rules of exclusion, it does not even require saying that any particular facts may not be considered.

At the same time, playing with proof levels is a blunt force tactic, one that can easily generate as many problems as it solves. Levels of proof that deviate farthest from the preponderance standard will generally be most effective in influencing trial court behavior and facilitating appellate monitoring, but they will also, at least in general, also generate more new bad outcomes.\textsuperscript{86} On the other hand, less severe adjustments will tend to have a lesser impact on trial court decisionmaking and make it harder and more costly for reviewing courts to monitor compliance. And to make matters worse, because the overall impact of any particular adjustment will depend on both the number of trial

\textsuperscript{82} See In re Winship, 397 U.S. 358, 361-64 (1970).
\textsuperscript{84} See, e.g., Fitzgerald v. Racing Ass’n of Cent. Iowa, 539 U.S. 103 (2003) (9-0 decision reversing a state supreme court’s invalidation of a state tax law under rational basis review).
\textsuperscript{85} Berman, \textit{supra} note 1, at 140. In general, “[i]f the goal is to minimize the number of erroneously decided cases . . ., the preponderance-of-the-evidence rule emerges as the superior choice.” Neil Orloff & Jerry Stedinger, \textit{A Framework for Evaluating the Preponderance-of-the-Evidence Standard}, 131 U. Pa. L. Rev. 1159, 1172 (1983). But this reasoning, Professor Berman has explained, “crucially depends on the assumption that the factfinder’s degree of subjective confidence regarding a given factual proposition . . . accurately corresponds to the statistical probability that that proposition is true.” Berman, \textit{supra} note 1, at 139. In contrast, if the Supreme Court has reason to believe that lower courts “systematically overestimate or underestimate the probative value of a particular type of evidence, or are otherwise systematically biased for or against a particular class or litigant,” \textit{id.} at 140, a heightened proof standard can reduce the total number of bad outcomes.
\textsuperscript{86} See \textit{supra} note 85.
court errors that were occurring before the change and the circumstances that produced those errors in the first place, there can be serious informational barriers to setting the optimal proof level.

All of the strategies discussed so far – rules, rules of exclusion, duties of explanation, baseline presumptions, and adjustments to the level of proof – directly regulate proceedings in trial courts. In so doing, they seek both to reduce the initial number of trial court errors and to ease their detection by appellate courts. But there is another family of techniques the Supreme Court may use when it seeks to increase its control over lower court outcomes: those that aim to eliminate barriers to effective appellate monitoring.

B. Increase Appellate Oversight

Trial courts do not always have the last word, but a number of doctrines frequently give them “a right to be wrong without incurring reversal.” Accordingly, another strategy the Supreme Court may use to control lower court outcomes is to limit what Professor Rosenberg called “secondary” discretion – “the degree of finality and authority a lower court’s decision enjoys in the higher courts.”

A certain amount of secondary discretion is both inevitable given present institutional arrangements and desirable on its own terms. It is inevitable because the ratio of trial to appellate judges makes it impossible for appellate courts – much less the Supreme Court itself – to engage in full-scale monitoring of every trial court ruling. And it is desirable because monitoring is costly and time-consuming for judges and litigants alike and because even the most self-confident Justice would probably agree that there are certain kinds of decisions that trial courts are more institutionally well-qualified to make.

The ultimate form of secondary discretion would be to make all trial court rulings entirely unappealable. But though the Supreme Court has long stated and recently reaffirmed that there is no freestanding constitutional right to an appeal, neither the Federal Government nor any State has gone that far.

---

87 Rosenberg, supra note 46, at 637.
88 Id.
89 See supra notes 32-35 and accompanying text.
90 See supra notes 32-35 and accompanying text.
What they do, however, is employ a number of doctrines that shield many trial court rulings from effective appellate scrutiny. For my purposes here, the three most important sources of secondary discretion are the final judgment rule, the adversity requirement, and deferential standards of appellate review.93

The first two doctrines are closely related. The final judgment rule, which is used in the federal system and most States,94 generally bars an appeal until the trial court has entered a definitive ruling in favor of one party, such as an award of damages to the plaintiff or a dismissal of the entire case.95 Because a great many cases settle or are voluntarily dismissed before that point, a huge number of trial court rulings – including those denying motions to dismiss or for summary judgment – are never subject to appellate review at all.96 The adversity requirement, in turn, bars appellate review unless the party seeking to challenge a particular ruling has suffered some legally cognizable consequence as a result of it.97 In situations where the final judgment rule is followed, this precept is generally interpreted to bar an appeal unless the party on the losing end of the challenged ruling was also on the losing side at the end of the litigation,98 meaning that some demonstrably incorrect lower court rulings are not reviewable even in theory.

Deferential standards of appellate review, in contrast, come into play once an appeal is successfully taken. Their effect – indeed, their purpose – is to give trial courts a degree of interpretive leeway. The well-settled principle that a trial court’s findings of historical fact will not be disturbed absent “clear error”99 and various doctrines stating that other sorts of decisions will be overturned...
only in the event of an “abuse of discretion” both constitute deferential standards of appellate review.

Because these doctrines can pose significant barriers to effective monitoring, it should be unsurprising that a Justice seeking to gain greater control over lower court outcomes may seek to modify them (assuming, that is, that the Justice thinks the intermediate appellate courts are, on balance, more likely to reach a good result than any given trial court judge). Accordingly, a sixth strategy the Supreme Court may use is to create exceptions to the final judgment rule, such as when it held that rulings denying motions to dismiss based on sovereign immunity are immediately appealable.

The upsides of such a strategy are obvious. The main one is lack of informational barriers: Like duties of explanation, simply authorizing more appeals requires no ex ante determination of whether certain facts should be dispositive (rules) or irrelevant (rules of exclusion), or even whether and to what extent the inquiry should be stacked in favor of one particular result (baseline presumptions and adjusted levels of proof). In addition, strategies that rely on expanded appellate monitoring should have little impact on incentives for non-judicial actors, and they are less to subject to manipulation by trial judges than those that seek to contract the scope of a trial court’s primary discretion.

These sizeable advantages, however, come with particular severe downsides. Unsuccessful appeals impose deadweight costs on the litigants and the judicial system, and increasing the overall number of appeals means reviewing courts will have less time to spend on each particular case. Even if they are comparatively more likely to reach the correct result, intermediate appellate courts will sometimes overturn trial court rulings that a majority of Justices would have regarded as correct. Creating exceptions to the final judgment rule also raises problems of legitimacy, both because the statutes governing federal court jurisdiction specify very few exceptions and because it would difficult to identify any source of authority for the Supreme Court to prescribe rules governing the timing of appeals in state courts. Finally,

---


102 See, e.g., 28 U.S.C. § 1292(a) & (b); Fed. R. Civ. Proc. 23(f).

103 Cf. Johnson v. Fankell, 520 U.S. 911 (1997) (declining to require state courts to permit interlocutory appeals even in situation where the underlying action was based on a federal statute and where an interlocutory appeal would have been permitted had case been litigated in federal court).
assuming the Court is not prepared to eliminate the final judgment rule entirely (a step that would be hugely costly and raise massive legitimacy problems), it will need to expend further time and effort devising ways to decide which interlocutory orders are nonetheless appealable and then enforcing those inevitably over- and under-inclusive restrictions against litigants’ attempts to evade them.104

Of course, having more appeals is likely to be of limited value if the rules governing them require substantial (or even complete) deference to the trial court. For that reason, a seventh strategy for increasing Supreme Court control over lower court outcomes is to lower the standard of appellate review. When the Supreme Court has held that various “constitutional facts” are subject to “independent examination” on appeal,106 it was employing this technique.

Whereas the benefits of increasing appellate scrutiny are basically the same as those of expanding the number of appeals, the downsides are somewhat different. Neither legitimacy nor implementation is likely to be a serious problem in most cases, because it is relatively rare for either the Constitution or federal statutes to specify a standard of review and because the standard’s application is within the control of the appellate courts themselves.107 Instead, the biggest problems with this approach are likely to be costs of enforcement (because truly non-deferential review can be quite time consuming for appellate courts) and creation of new legal errors (because even time-strapped trial courts may simply be better at performing certain kinds of tasks than time-strapped appellate panels).

There are times, however, when neither increasing the availability of appeals nor intensifying their focus will be enough to overcome the difficulties an appellate court can face in attempting to understand precisely what transpired before a different judge in another courtroom. In those situations, the Supreme Court can employ an eighth and final strategy for increasing its

104 See, e.g., Digital Equip. Corp. v. Desktop Direct, Inc., 511 U.S. 863, 868 (1994) (noting that applicability of the “collateral order” doctrine “is to be determined for the entire category to which a claim belongs, without regard to the chance that the litigation at hand might be speeded or a particular injustice averted by a prompt appellate decision.” (internal quotation marks and citation omitted)).

105 According to Professor Charles Alan Wright, the prevailing rule was once that reviewing courts had no power whatsoever to review a trial court’s refusal to grant a new trial on the theory that the verdict was against the weight of the evidence. See The Doubtful Omniscience of Appellate Courts, 41 MINN. L. REV. 751, 758-60 (1957).


107 See Rosenberg, supra note 46, at 641. But see infra note 151.
control over lower court outcomes: it can use (and direct intermediate appellate courts to apply) various appellate presumptions.

An appellate presumption is simply a baseline presumption directed at appellate courts rather than trial courts. To return once again to *North Carolina v. Pearce*:\(^{108}\) When the Supreme Court held that a trial court’s failure to state on the record its reasons for imposing a longer sentence following a second trial raises a presumption of vindictiveness on appeal,\(^ {109}\) it was announcing an appellate presumption.

As *Pearce* illustrates, one way appellate presumptions can increase Supreme Court control is by helping ensure trial court compliance with duties of explanation. But their usefulness expands beyond that relatively narrow compass. Indeed, appellate presumptions can be used to enforce nearly any restriction on a trial court’s primary discretion, as illustrated by the Court’s holding in *Chapman v. California*\(^ {110}\) that any uncorrected constitutional error during a criminal trial must be presumed prejudicial on appeal.\(^ {111}\) Were *Chapman*’s holding firmly applied and consistently followed,\(^ {112}\) it would both simplify a reviewing court’s task and give trial courts an additional incentive to get it right the first time.

At least in the criminal context, which is where both *Pearce* and *Chapman* hold sway, the chief downside of appellate presumptions is that they can be extremely costly to enforce. Judicial and litigant resources are wasted every time a reviewing court requires further proceedings simply because a trial court failed to say the right thing or based on an error that had no impact on the outcome. To make matters worse, appellate presumptions may also lead to increased costs for appellate courts by encouraging litigants who have suffered

---


\(^{109}\) See id. at 726.

\(^{110}\) 386 U.S. 18 (1967).

\(^{111}\) See id. at 24. The opposite presumption applies when a federal court is entertaining a petition for a writ of habeas corpus. Under *Brecht v. Abrahamson*, 507 U.S. 619 (1993), a habeas petitioner cannot obtain relief based on even a conceded constitutional violation unless she “can establish that [the violation] resulted in ‘actual prejudice.’” Id. at 638; see also *Fry v. Piler*, 127 S. Ct. 2321 (2007) (holding that *Brecht* applies in situations where the state appellate courts did not recognize the constitutional error and thus did not apply *Chapman*-style harmless-error review).

no concrete prejudice to nonetheless take an appeal. 113 Perhaps for these reasons, the Supreme Court has shown itself quite reluctant in recent years to adopt any such presumptions, either with respect trial court violations of duties of explanation 114 or more generally. 115

* * *

As Judge Jerome Frank observed more than sixty years ago, trial courts will always have one inestimable advantage in any struggle with their judicial superiors: the ability to find (and thus characterize) the underlying facts, findings that for reasons of both necessity and sound practice will almost always be accorded great deference on appeal. 116 Given that reality, it is unsurprising that many of the techniques discussed in this Part can be thought of, at least in part, as attempts to shrink that advantage. Rules restrict the number of facts that are potentially significant. Rules of exclusion bar reliance on certain types of facts whose presence or absence can be particularly difficult to monitor. Presumptions and raising the level of proof change the strength or type of facts required to justify a disfavored result. Duties to explain and appellate presumptions attempt to force trial judges to identify the facts on which they are actually relying, rather than requiring reviewing courts to rule out all bases on which they could possibly be relying. Expanded appellate review facilitates at least some scrutiny of more trial court determinations, including factual ones. And lowering the standard of appellate review lessens the deference owed to certain determinations that may otherwise be seen as “factual” in nature.

At this point, it seems appropriate to move from possibilities to actualities. Part I explained why the Supreme Court has reason to worry about lower court compliance with its decisions. This Part argued that there are a number of ways in which the Justices can craft legal doctrines to address that risk. The next and final Part will identify a number of discrete areas in which the modern Court appears to have done just that.

113 See Roger Traynor, The Riddle of Harmless Error 50 (1970) (“Reversal for error, regardless of its effect on the judgment, encourages litigants to abuse the judicial process . . . .”).


III. THE STRATEGIES EMPLOYED

The claim that the Supreme Court has sometimes crafted legal doctrines based on mistrust of lower courts should be largely uncontroversial. Distrust of state courts, particularly southern state courts, is generally viewed as an important theme of the Warren Court’s jurisprudence.\textsuperscript{117}

But distrust of lower courts did not end with the Warren era Justices and it need not result in doctrines likely to please political liberals. To the contrary, this Part argues that there are at least three areas in which the Supreme Court’s recent actions suggest an attempt to exert greater control over all lower courts and to nudge them in directions more consistent with political conservatism. Not only do these developments suggest that concerns about lower courts are alive and well, they also underscore that matters are far more complicated than the familiar debates regarding rules versus standards and provide concrete illustrations of the relationships between various strategies for increasing Supreme Court control over lower court decisionmaking.

Although my three examples are somewhat eclectic, there are reasons for each choice. The first – the constitutional law of punitive damages – supplies the clearest narrative of increasing frustration with lower courts as a motivator for doctrinal development and suggests that concerns about state courts are alive and well nearly four decades after Earl Warren relinquished his seat. The second example – the rules governing “officer suits” brought under 42 U.S.C. § 1983 – shows that the Supreme Court is perfectly capable of fashioning doctrine based on mistrust of lower federal courts and of doing so when interpreting statutes as well as the Constitution. The final example – the concept of “reasonable” searches and seizures under the Fourth Amendment – provides the most direct contrast with the Warren Court and makes clear that a desire to rein in lower courts can just as easily afflict a Court looking to narrow federal rights as to expand them.

Two clarifications seem warranted at the outset. First, this Part’s approach is largely inductive. Ascribing motivations to a multi-member body is difficult at best,\textsuperscript{118} and it is made more so here by the fact that norms of judicial behavior would seem to prevent Supreme Court Justices from announcing that they are acting in a particular way based, even in part, on distrust of lower court judges. Accordingly, this Part must make the comparatively more limited claim that lower court distrust is one fairly compelling explanation for an


\textsuperscript{118} Cf. Frank Easterbrook, Statutes’ Domains, 50 U. CHI. L. REV. 533, 547 (1983) (suggesting that though “[e]ach member may or may not have a design,” a multi-member group such as a legislature “has only outcomes”).
overall course of decisions. Second, all of the decisions discussed in this Part could be criticized on one ground or another. Because this Article is largely descriptive, it takes no position regarding proper interpretive methods, nor does it seek to defend the legal correctness or substantive wisdom of any of these decisions.

A. Modern Distrust of State Courts: The Constitutional Law of Punitive Damages

Punitive damages awards have long been subject to some sort of judicial review for “reasonableness” in most United States jurisdictions. But it is only recently that the Supreme Court has held that the size of every such award is subject to federal constitutional restrictions. More so than in perhaps any other area, the development of the Supreme Court’s punitive damages jurisprudence during the last several decades suggests an increasing frustration with lower courts in general and state judiciaries in particular.

Although the Court has long asserted the power to control punitive damages awards in cases brought under federal causes of action, it has been far slower to do so when the underlying dispute is governed by state law. Indeed, as recently as 1989, a seven-Justice majority wrote that in situations “where state law provides the basis of decision, the propriety of an award of punitive damages for the conduct in question and the factors the jury may consider in determining their amount, are questions of state law.”

Notwithstanding that statement, the Court had actually begun asserting some modest control over state law punitive damages awards more than two decades earlier. In two decisions issued during the mid-1960s, it construed the federal labor laws as implicitly forbidding state law-based punitive damages in certain situations. And in 1974, it held that the First Amendment creates

---

special restrictions on punitive damages awards in state law defamation actions. But despite saying some fairly critical things about punitive damages in general, these decisions were sharply limited in scope, tied as they were to particularized statutory or constitutional commands.

Starting in the 1980s, however, various opinions began intimating that the Federal Constitution may impose more general restraints on state court punitive damages awards. In 1986, a majority opinion that had already resolved the case on other grounds went out of its way to describe a party’s argument that both the Due Process and Excessive Fines Clauses limited state law punitive damages awards as raising “important issues which, in an appropriate setting, must be resolved.” Two years later, the Court granted review to consider those issues, but failed to reach them after concluding that the appealing party had neglected to raise them with the state courts. In so doing, however, Justice Marshall’s majority opinion described the Excessive Fines question as one “of some moment and difficulty,” and Justice O’Connor’s concurring opinion declared that “there [was] reason to think” that the Due Process Clause imposed restrictions on punitive damages.

The Court’s next three decisions were characterized by increasingly firm statements that the Federal Constitution imposed at least some limits coupled with an apparent reluctance on the part of a majority of Justices to identify precisely what they might be. In 1989, the Court issued an opinion that pointedly noted that the parties agreed “that due process imposes some limits on jury awards of punitive damages,” but nonetheless declined to say more because of the way in which the parties had litigated the case. In a separate opinion, Justice O’Connor claimed that “[a]wards of punitive damages are skyrocketing,” criticized state appellate courts for sustaining punitive awards far higher than those upheld in even the recent past, and stressed that “nothing in the Court’s opinion forecloses a due process challenge to awards of punitive damages or the method by which they are imposed.”

---

124 See id. at 350 (“[J]uries assess punitive damages in wholly unpredictable amounts bearing no necessary relation to the actual harm caused”).
127 Id. at 79.
128 Id. at 87-88 (O’Connor, J., concurring in part and concurring in the judgment).
130 Id. at 282-83 (O’Connor, J., concurring in part and dissenting in part).
Next came 1991’s *Pacific Mutual Life Insurance Co. v. Haslip*, where a six
Justice majority finally found a case where they deemed the issues properly
preserved and rendered an opinion that seemed premised on the view that the
Due Process Clause imposed both substantive and procedural restraints on
 punitive damages awards. But although the *Haslip* Court stated that the
 punitive award before it “may be close to the [constitutional] line,” it actually
 rejected five specific due process attacks, including the claim that punitive
damages should be unavailable absent “clear and convincing” proof of
wrongdoing. In addition, the Court stated that it “need not, and indeed . . .
cannot, draw a mathematical bright line between the constitutionally
acceptable and the constitutionally unacceptable that would fit every case.”
In dissent, Justice O’Connor argued that the due process issues raised by
 punitive damages were “ripe for reevaluation,” in large measure because
changes in state law made them available in more types of cases and because
juries were returning, and appellate courts were sustaining, far higher awards
than they had in the past.

The apogee of the Court’s “talk tough, do little” phase came in 1993’s
*TXO Production Corp. v. Alliance Resources Corp.* For the first time, a clear
majority stated that a “grossly excessive” punitive award would “violate the
substantive component of the Due Process Clause.” At the same time, the
Court declined to upset a punitive award that was 526 times the amount of the
compensatory damages, and its language seemed designed to preclude any
serious attempt to exert Supreme Court control over punitive damages.
Repeating *Haslip’s* rejection of clear rules, Justice Stevens’ plurality opinion
stressed that “a jury imposing a punitive damages award must make a
qualitative assessment based on a host of factors and circumstances unique to
the particular case before it” and added that any award that was the product of

---

132 Id. at 9, 23-24.
133 Id. at 24.
134 Id. at 23 n.11; see also id. at 18 (“[W]e cannot say that the common-law method for
assessing punitive damages is so inherently unfair as to deny due process and be per se
unconstitutional”); id. at 14 (rejecting argument that it violates due process to make
corporation vicariously liable in punitive damages for unauthorized actions by its employees);
id. at 19-22 (rejecting argument that state-court approved jury instructions gave too much
discretion to the factfinder); id. at 23-24 (rejecting due process attack on the size of the award).
135 Id. at 18.
136 Id. at 61-63 (O’Connor, J., dissenting).
138 Id. at 458 (plurality opinion); see id. at 479-86 (O’Connor, J., dissenting).
139 See id. at 453.
“fair procedures . . . is entitled to a strong presumption of validity.”\textsuperscript{140} The plurality also declined to adopt two proposed rules of exclusion, describing it as “well-settled” that both a defendant’s wealth and alleged wrongdoing in other parts of the country may be considered both by juries in deciding whether to impose punitive damages and by lower courts in deciding whether to uphold such awards.\textsuperscript{141}

The first Supreme Court decision to vacate a state court punitive damages award on due process grounds is not terribly well known. And although the five million dollar punitive award in that case was more than six times the already substantial compensatory damages,\textsuperscript{142} the basis for the Supreme Court’s action was not that the punitive figure was necessarily too high; nor did the Court hold that it had been based on improper information or imposed pursuant to an improper standard of proof. Rather, 1994’s \textit{Honda Motor Co. v. Oberg}\textsuperscript{143} was a case about constitutionally compelled \textit{monitoring}. It declared unconstitutional a provision of the Oregon state constitution that had been interpreted to bar all judicial review of the size of a jury’s punitive damages award.\textsuperscript{144}

Because Oregon was apparently the only State with this sort of rule, \textit{Oberg}’s impact was necessarily quite limited. The same cannot be said, however, of the decision that first made plain there had been a real shift in the Supreme Court’s approach – its 1996 ruling in \textit{BMW of North America, Inc. v. Gore}.\textsuperscript{145}

\textit{Gore} is significant for two main reasons. One, it was the first time the Supreme Court held that the sheer size of a punitive damages award exceeded constitutional limits. Indeed, the Court did so despite the fact that the Alabama state courts had already reviewed the award using procedures whose constitutionality the Court had blessed in \textit{Haslip} and cut the size of the jury’s award in half.\textsuperscript{146} Two, although Justice Stevens’ majority opinion declared that the Court was still “not prepared to draw a bright line marking the limits of a constitutionally acceptable punitive damage award,”\textsuperscript{147} it laid down three “guideposts” for assessing constitutionality, and gave extensive (albeit, largely case-specific) guidance for their application.\textsuperscript{148}

\begin{itemize}
  \item[140] \textit{Id.} at 457 (plurality opinion).
  \item[141] \textit{Id.} at 462 n.28.
  \item[143] 512 U.S. 415 (1994).
  \item[144] \textit{See id.} at 426-35.
  \item[145] 517 U.S. 559 (1996)
  \item[146] \textit{Id.} at 566-67.
  \item[147] \textit{Id.} at 585.
  \item[148] \textit{Id.} at 574-85.
\end{itemize}
Although the Court did not consider the substantive constitutionality of another punitive damages award for a number of years after *Gore*, the pattern of increasing constitutional scrutiny of trial court punitive damages rulings continued unabated. In 2001, *Cooper Industries, Inc. v. Leatherman Tool Group, Inc.*[^149] held that a United States Court of Appeals had erred in deferring to a trial court’s application of the three *Gore* “guideposts.”[^150] In doing so, the Court brushed aside arguments that the appropriate size of a punitive damages award presents an issue of “fact” that the Federal Constitution’s Re-Examination Clause shields from scrutiny by a federal appellate court[^151], and described *de novo* appellate review as necessary to permit “appellate courts to maintain control of, and to clarify, the legal principles.”[^152] In short, whereas *Gore* focused on limiting a trial court’s primary discretion in entering judgment on a jury’s punitive damages award and did so by making the underlying constitutional analysis somewhat more rule-like, *Cooper Industries* aimed to reduce the level of secondary discretion created by mandating a non-deferential standard of appellate review.

Then, in 2003, the Supreme Court issued an opinion that seemed to embody multiple strategies for controlling state court punitive damages awards. Although Justice Kennedy’s majority opinion in *State Farm Mutual Automobile Insurance Co. v. Campell*[^153] “decline[d] again to impose a bright-line ratio which a punitive damages award cannot exceed,” it appeared to move the underlying constitutional test substantially further in the direction of a rule both by pointedly suggesting that ratios above 9:1 are presumptively suspect[^154], and generally discussing each “guidepost” in such detail that Justice Ginsburg accused the majority of converting them into “marching orders” for the lower courts.[^155] *State Farm* also appeared to endorse at least two rules of exclusion, stating that courts may not seek to justify punitive damages awards based on

[^150]: See id. at 436.
[^151]: See id. at 437-39. The Re-examination Clause provides that: “[N]o fact tried by a jury, shall otherwise be re-examined in any Court of the United States, than according to the rules of the common law.” U.S. CONST. amend. VII. Although this provision’s text does not speak to the relationship between different courts, the Supreme Court has held that it incorporates Founding-era understanding that trial courts may disturb jury verdicts in certain situations where appellate courts would not be permitted to do so. See Gasperini v. Center for Humanities, Inc. 518 U.S. 415, 432-34 (1996).
[^152]: See *State Farm*, 532 U.S. at 436 (quoting *Ornelas v. United States*, 517 U.S. 690, 697 (1996)) (emphasis added).
[^154]: See id. at 425 (“[F]ew awards exceeding a single-digit ratio between punitive and compensatory damages, to a significant degree, will satisfy due process.”).
[^155]: Id. at 439 (Ginsburg, J., dissenting).
“unrelated” out-of-state conduct or a defendant’s wealth. In sharp contrast to the Court’s earlier declaration in TXO, State Farm also suggested that the baseline presumption was that any punitive damages award – even one imposed pursuant to scrupulously fair procedures – is constitutionally suspect. And finally, without any acknowledgment that the earlier case had been litigated in federal court or that the Due Process Clause has been construed so as to confer no right to appellate review in the first place, State Farm flatly stated that Cooper Industries had “mandated” that all appellate courts “conduct de novo review of a trial court’s application of [the Gore factors] to [a] jury’s award.”

The Court’s decision last Term in Philip Morris USA v. Williams is interesting for two reasons. The first is its actual holding: the Court adopted another rule of exclusion, declaring that States may not use punitive damages awards to punish (and courts may not seek to justify the constitutionality of a punitive damages award by making reference to) harms to third parties who are not before the court. But in some ways even more interesting is Part II of Justice Breyer’s majority opinion, which took pains to reassert, in almost bullet form, the main holdings of Cooper Industries, Oberg, Gore, and State Farm.

It may be possible to tell several stories about the development of the Supreme Court’s punitive damages jurisprudence. A fairly compelling one,

---

156 See id. at 421.
157 See id. at 427.
158 See id. at 419 (“It should be presumed a plaintiff has been made whole for his injuries by compensatory damages, so punitive damages should only be awarded if the defendant’s culpability, after having paid compensatory damages, is so reprehensible as to warrant the imposition of further sanctions to achieve punishment or deterrence.”).
160 See State Farm, 538 U.S. at 418.
162 See id. at 1063-64.
163 See id. at 1062-63.
164 See, e.g., Samuel Issacharoff & Catherine M. Sharkey, Backdoor Federalization, 53 UCLA L. REV. 1353, 1356 (2006) (viewing constitutionalization of punitive damages as part of broader pattern of Supreme Court attempts “to capture the considerable benefits that flow from national regulatory uniformity and to protect an increasingly unified national (and international) commercial market from the imposition of externalities by unfriendly state legislatures”); Andrew M. Siegal The Court Against the Courts: Hostility to Litigation as an Organizing Theme in the Rehnquist’s Court’s Jurisprudence, 84 TEX. L. REV. 1097, 1108, 1146-52 (2006) (using punitive damages example to support claim that “it is impossible to understand the [Rehnquist] Court’s complicated intellectual matrix without acknowledging and assimilating the Court’s hostility towards the institution of litigation and its concomitant skepticism as the ability of litigation to function as a mechanism for organizing social relations and collectively administering justice”).
however, would involve an increasing conviction that juries had gotten out of control—and that at least some state courts had defaulted on their obligation to reign them in—gradually overcoming whatever reluctance various Justices otherwise would have had about federalizing almost quintessential issues of state law. Starting with the Court’s 1938 abandonment of the general law in *Erie Railroad Co. v. Tompkins*[^165] and through the early years of what some perceived as “revolutions” in consumer and mass tort litigation[^166], the Justices were apparently content to leave regulation of punitive damages almost entirely to the state judiciaries, though they made exceptions for two areas of special federal concern (labor and libel). Then, starting with Justice O’Connor—herself a former state court judge[^167] and normally a champion of judicial federalism[^168]—various Justices became increasingly concerned that at least some state courts were not holding up their end of the bargain[^169], though Court majorities initially declined to do much else[^170]. By the mid-1990s, a majority of Justices were finally willing to take action, though their initial efforts involved invalidating one extreme example of judicial abdication (*Oberg*) and overturning one outlier award in an opinion that hewed close to the facts of the particular case (*Gore*). By the early years of this century, however, the Court had grown bolder and/or more fed up, as demonstrated by its substantial tightening of the *Gore* “guideposts” (*State Farm*) and how the requirement that reviewing courts undertake some scrutiny of punitive damages awards morphed into a command to performing exacting appellate review (*Cooper Industries*).

At the same time, developments in this area also nicely illustrate the trade-offs of, and the relationships between, various strategies for controlling lower

[^165]: 304 U.S. 64 (1938). Before *Erie*, the Court had reviewed the permissibility of punitive damages awards in diversity cases litigated in federal court. See, e.g., Lake Shore & M.S. Ry. Co. v. Prentice, 147 U.S. 101 (1893) (setting forth standards under which railroad could be held liable in punitive damages based on behavior of conductor towards passenger).


[^168]: See, e.g., Coleman v. Thompson, 501 U.S. 722, 726 (1991) (O’Connor, J.) (beginning opinion for the Court in a capital case with the words: “This is a case about federalism. It concerns the respect that federal courts owe the States and the States’ procedural rules . . . .”)


[^170]: The epitome may be the following statement from the *Haslip* majority opinion: “We note once again our concern about punitive damages that ‘run wild.’” Pac. Mut. Life Ins. Co. v. Haslip, 499 U.S. 1, 18 (1991). The quotation marks are Justice Blackmun’s.
court behavior. Although the Court’s doctrines have certainly grown much more determinative during the last 15 years, as recently as 2003 it declined to impose a firm bright-line rule, whether because a majority of Justices have decided that hard rules are impractical in this area\textsuperscript{171} or because there is no majority for any given one. Unable to coalesce around a simple and easily police-able requirement, the evidence suggests that the Justices have increasingly employed various other strategies, including rules of exclusion, shifting the baseline presumption, and expanding appellate review.

If vigorously implemented by appellate courts, the present system may be adequate to ensure meaningful control over juries and trial courts. Punitive damages are rarely entered until at or shortly before the time of final judgment. The stakes involved in such cases are likely to make an appeal worthwhile. And the well-settled rule that a reviewing court lacks the power to increase a punitive damages award\textsuperscript{172} means that the risks of fighting such awards will usually point in one direction.\textsuperscript{173} When all of these factors are coupled with the prospect of \textit{de novo} appellate review, defendants’ incentives to take appeals are quite high.

But all of this starts with a pretty big “if.” Indeed, the primary downside of the Court’s current approach is that it rests almost entirely on the willingness and ability of \textit{state appellate courts} to serve as proxies for the Supreme Court itself. In situations where the plaintiff’s cause of action is created by federal law, either party may generally decide to have the case heard and decided by the lower federal courts,\textsuperscript{174} and it has long been argued that one of the primary reasons for the development of the “modern” post-conviction writ of habeas corpus was to permit the Supreme Court to enlist the lower federal courts to help ensure compliance with its revolutionary criminal procedure decisions.\textsuperscript{175} In contrast, punitive damages are most often awarded in cases governed by state law and in situations where ordinary preclusion principles would preclude a “collateral” attack in a new federal court proceeding.\textsuperscript{176} Accordingly, the

\textsuperscript{171} Cf. Mathias v. Accor Econ. Lodging, Inc. 347 F.3d 672, 678 (7th Cir. 2003) (Posner, J.) (“The judicial function is to police a range, not a point.”).


\textsuperscript{173} If a trial court has reduced a jury’s punitive damages award, though not as much as the defendant would have liked, filing an appeal does carry some risk, though in that case it is likely that the plaintiff will appeal even if the defendant does not do so.

\textsuperscript{174} See 28 U.S.C. §§ 1331 (original jurisdiction); 1441(a) (removal jurisdiction).

\textsuperscript{175} See, e.g., Barry Friedman, \textit{A Tale of Two Habeas}, 73 MINN. L. REV. 247, 273-77 (1988).

\textsuperscript{176} Cf. Pennzoil Co. v. Texaco, Inc., 481 U.S. 1 (1987) (holding that federal court should have abstained from considering constitutional attack on the largest punitive damages award ever rendered as of that date).
Supreme Court is the only federal court authorized to consider the constitutionality of most punitive damages awards, and the highly fact-dependent nature of even the post-State Farm “guideposts” and the massive constraints on their time to review certiorari petitions mean that there is little the Justices can do under the present regime to ensure appellate court compliance other than throw out an occasional outlier award.

So what will happen if a majority of Justices continue to perceive that state courts have been insufficiently responsive to their aim to rein in punitive damages awards? At some point, the Justices may find themselves sufficiently informed to formulate and willing to bear the costs of implementing a rule, and either impose an absolute or presumptive ratio with respect to compensatory damages that punitive awards may not exceed or hold that punitive damages are categorically unavailable in certain circumstances or absent some specified findings. The Court could also squarely hold that punitive damages are presumptively unavailable as a constitutional matter, or revisit another one of its earlier rulings and hold that punitive damages may not be imposed absent clear and convincing evidence of sufficiently culpable conduct. Having mandated de novo appellate review of punitive damages awards, the Court could conceivably require special verdict forms so that reviewing courts could have the benefit of the jury’s actual basis for decision rather than a trial court’s attempted reconstruction, or hold that the Constitution itself requires a stay of execution of judgment until any appellate review is complete.

Regardless of what happens going forward, however, the events that have already transpired illustrate the usefulness of thinking about distrust of lower courts as a motivating factor in the creation of Supreme Court doctrine. And as the next Section demonstrates, the distrust need not be limited to state judiciaries or play out in the constitutional arena.

B. Reining in Federal Trials Courts: Constitutional Tort Suits

The previous Section told a fairly familiar story from the Warren era – increasing distrust of state courts leading to constitutionalization of matters traditionally governed by state law – about far more recent developments in the punitive damages context. But though it has often been believed that the


178 For a dated but still marvelous illustration of just how little time the Justices have, see Henry M. Hart, Jr., The Supreme Court, 1958 Term—Foreword: The Time Chart of the Justices, 73 HARV. L. REV. 84 (1959).

lower federal courts are generally better agents of the Supreme Court than their state counterparts, that does not mean the Justices have been willing to assume perfect compliance by the former. Indeed, as this Section explains, recent developments in at least one area may be best explained as efforts to control lower federal courts in general – and federal trial judges in particular.

Citizens who believe their constitutional rights have been violated have long been able to sue for money damages. Although sovereign immunity shields the government itself from suits brought by individuals, it is well-established that there is generally no sovereign immunity problem with suing the appropriate government official in her “personal” capacity, even for actions performed on the government’s behalf. These actions are generally known as “officer suits.”

For much of our Nation’s history, well-settled jurisdictional principles meant that suits seeking money damages from state officials almost invariably had to be litigated in state court under state-supplied rules. The plaintiff would accuse the defendant of committing some traditional common law wrong, such as trespass, assault, or false imprisonment. The defendant would answer by claiming “official privilege” – that is, government authorization – as a defense. The plaintiff would respond by invoking the superior authority of the Federal Constitution and arguing that it stripped the defendant of her immunity, leaving her indistinguishable from any other law-breaker. But because the issue of federal law was not a necessary ingredient of the plaintiff’s “well-pleaded complaint,” such suits could neither be filed in federal court in the first instance nor removed there by the defendant. It also meant the plaintiff was required to fit her suit into a state-law recognized cause of action, and to contend with any restrictions state law placed on appropriate remedies.

---

180 See, e.g., Neuborne, supra note 39, at 1124-25.
181 See Hans v. Louisiana, 134 U.S. 1 (1890).
183 In contrast, actions filed against federal officers have long been removable. See HART & WECHSLER, supra note 19, at 905.
185 See generally 28 U.S.C. § 1441(a) (setting out general rule that case may be removed only if action could originally have been filed in federal court). In fact, until Congress enacted § 1441(a)’s predecessor in 1875, “no general grant of removal jurisdiction in ‘arising under’ cases existed.” HART & WECHSLER, supra note 19, at 905.
All that changed in 1961 when the Supreme Court dusted off an obscure federal statute,187 and held that it authorized federal court litigation under federally supplied rules in virtually every situation where a citizen accuses a state official of violating her constitutional rights. Enacted in 1871, 42 U.S.C. § 1983 authorizes “an action against law, suit in equity, or other proper proceeding for redress” against any person who violates another’s constitutional rights while acting “under color of any statute, ordinance, regulation, custom, or usage, of any State or Territory.” The principal question in Monroe v. Pape188 was whether state officers whose conduct had violated the constitution and laws of Illinois had nonetheless acted “under color of” that State’s law for purposes of § 1983, notwithstanding the fact that the state courts stood able and willing to provide redress.189

By an 8-1 vote, the Supreme Court said “yes.” And although Justice Douglas’ majority opinion is famously difficult to follow,190 its reasoning appears to rest in large measure on distrust of state court judges. Canvassing § 1983’s legislative history, Justice Douglas reasoned that the statute aimed not only to “override certain kinds of state laws” and to “provide[] a remedy where state law was inadequate,” but also “to provide a federal remedy where the state remedy, though adequate in theory, was not available in practice.”191 But because Monroe’s holding was that the federal remedy is always available,192 this reasoning appears to make sense only if one posits a profound and continuing distrust of state courts’ willingness to vindicate federal constitutional rights and a corresponding desire to enable such suits to be brought in the presumably more trustworthy federal courts.193

That the birth of modern § 1983 was founded in large measure on mistrust of state court judges is fairly well-recognized.194 What seems to have largely

---

187 One commentator found only 21 reported cases brought under the statute between its enactment in 1871 and 1920. See Comment, The Civil Rights Act: Emergence of an Adequate Federal Civil Remedy?, 26 Ind. L.J. 361. 363 (1951).
189 See id. at 172.
192 See id. at 183.
193 Another possible justification for Monroe’s holding – though not one articulated by Justice Douglas – involves the difficulty of making case-by-case assessments about whether state remedies are “adequate” and “available.” See JEFFRIES ET AL., supra note 190, at 46.
194 See, e.g., Neuborne, supra note 39, at 1110 (describing Monroe as resting on “thinly disguised assumptions of nonparity between state and federal courts”).
escaped commentary, however, are several ways in which post-
Monroe developments can be seen as motivated by concerns about federal trial judges.

Monroe largely removed state courts from the picture of what has come to be called “constitutional tort litigation.” Because the plaintiff’s cause of action is now supplied by a federal statute, these suits may always be filed in federal court. In addition, although the Supreme Court later clarified that § 1983 actions can be filed in state court as well – indeed, that state courts are generally required to entertain them – it is also clear that defendants retain the option to remove such suits to federal court. Barring a mistake, therefore, the Justices know the only way a § 1983 action will be litigated in state court is if both parties choose for it to be there. And given that, the only reason to employ any of the strategies described in Part II would seem to be concerns about the abilities and propensities of the lower federal courts.

Yet evidence of those strategies abounds in constitutional tort jurisprudence, most notably in the context of doctrines governing various “immunities.” Not long after Monroe reinvigorated § 1983, the Court began expressing anxiety that both ultimate damages liability and the prospect of being forced to defend against time-consuming and expensive lawsuits risked “over-detering” vigorous and socially desirable conduct by government officials. In response to these concerns, the Court developed various “immunity” defenses and has continually refined the principles governing their application over the succeeding decades.

The first set of questions involved the nature of the defenses and the identity of the officials entitled to claim them. For a small group of officials – the President of the United States, legislators, and judges – the Court settled on a bright-line rule: Absolute immunity for all acts even arguably taken in the relevant capacity. Although the Court has acknowledged that this approach will shield some culpable and socially undesirable conduct from legal consequence, it has stressed the existence of alternate control mechanisms and

---


the need give the relevant officials an especially wide field in which to operate free from judicial scrutiny.\footnote{See Bogan, 523 U.S. at 52-53; Nixon, 457 U.S. at 751-53, 757; Stump, 435 U.S. at 363-64.}

In the main, however, the nature of the \S\ 1983 context has made it difficult to rely on rules as a strategy for controlling lower court decisionmaking. Having interpreted \S\ 1983 to authorize suits against government officials for what the Justices presumably regarded as good and sufficient reasons, it would seem perverse to hold that those very same officials are all protected by an absolute immunity defense. More generally, \S\ 1983 is transsubstantive – that is, it authorizes suits seeking to vindicate a wide variety of rights against a diverse array of officials who act in an almost limitless variety of circumstances. Given that characteristic, the informational costs of creating any across-the-board rules would be massive and even the best possible ones would almost certainly be substantially over- and under-inclusive with respect to their underlying purpose of creating the socially optimal level of deterrence.

For these reasons, it seems unsurprising that the Supreme Court has generally chosen to articulate the scope of immunity defenses in terms of standards rather than rules. More than three decades ago, it made clear that the vast majority of government officials are eligible for what has come to be called “qualified immunity.”\footnote{See Scheuer, 416 U.S. at 247. Accord Harlow v. Fitzgerald, 457 U.S. 800, 807 (1982).} And although the Court’s precise articulation of the test has varied over the years, the classic one is a quintessential standard: Even government officials who have violated a plaintiff’s constitutional rights, the Court has stated, “are shielded from liability for civil damages insofar as their conduct [did] not violate clearly established statutory or constitutional rights of which a reasonable person would have known.”\footnote{Harlow, 457 U.S. at 818.}

Of course, the main downside of an approach that asks whether a given point of law was “clearly established” or what a hypothetical “reasonable person” should have known is that it gives a great deal of leeway to the initial decisionmaker – here, federal trial judges. That said, the Supreme Court has taken a number of steps that seem designed to facilitate effective monitoring of these decisions by the Courts of Appeal.

Most notably, the Supreme Court has carved out a massive exception to the final judgment rule. Orders denying motions for summary judgment are among the classic “interlocutory” orders generally shielded from appellate scrutiny until all trial court proceedings are complete.\footnote{See, e.g., 10A Charles Alan Wright et al., Federal Practice and Procedure \S\ 2715 (3d ed. 2007).} Yet in Mitchell v. Forsyth,\footnote{472 U.S. 511 (1985);} the Court held that orders denying a defendant’s request for...
summary judgment based on qualified immunity are immediately appealable. Not only does Mitchell substantially increase the odds these types of summary judgment denials will be subject to meaningful appellate scrutiny, the very existence of this unique review mechanism may encourage district courts to grant more of the underlying motions in the first place.

The Supreme Court has also taken a number of steps to ensure that the additional appellate monitoring authorized by Mitchell will be as effective as possible. For one thing, it has announced a lowered standard of appellate review. Notwithstanding the inherently fact-bound nature of whether the relevant law had been sufficiently established to give an officer “fair warning” that her particular conduct transgressed constitutional norms, the Court has been adamant in viewing the ultimate question as one of law that is reviewable de novo on appeal.

In addition, the Court has adopted an important rule of exclusion that facilitates more effective appellate monitoring. In Harlow v. Fitzgerald, the Court eliminated the requirement that a defendant seeking qualified immunity must establish that she acted with subjective good faith. In doing so, the Court later explained, Harlow rendered a defendant's state of mind “simply irrelevant to the qualified immunity defense.” Although the Court’s main stated justifications for this innovation were that subjective inquiries are intrusive and too often prevent “the defeat of insubstantial claims without resort to trial,” it may also be significant that disputes regarding a person’s mental state may be the situation in which trial courts’ informational advantage over reviewing courts is the greatest.

Finally, and perhaps most controversially, the Supreme Court has also imposed what can be seen as a duty of explanation on all lower courts. In a line of decisions culminating with 2001’s Saucier v. Katz, the Court has insisted that lower court confronting motions for summary judgment based on qualified immunity must analyze whether there has been a constitutional violation before proceeding to consider whether the defense has been made

---

207 457 U.S. 800 (1982).
208 See id. at 813-17.
210 Harlow, 457 U.S. at 813-14, 815-17.
out—a ruling that has been greeted with a pronounced lack of enthusiasm by many lower court judges.213

At the same time, developments in the constitutional tort arena also underscore the limitations of, and downsides to, an approach that relies predominantly on close appellate monitoring to secure trial court compliance. In *Johnson v. Jones*,214 the Supreme Court unanimously held that although a trial court’s *ultimate* conclusions that a constitutional violation occurred and that the defendant is not entitled to summary judgment based on qualified immunity are immediately appealable under *Mitchell*, the interpretations of fact on which those conclusions are based are not subject to appellate scrutiny.215 Though not denying that this holding will sometimes shield erroneous trial court rulings from appellate review,216 the Court cited, among other things, considerations of institutional competence and a desire to avoid consuming vast quantities of appellate court time on fact-dependent issues that may simply be resolved at trial.217

Although the Supreme Court has relied mostly on monitoring strategies in this context, it has not foreclosed entirely attempts to influence initial trial court decisions. Despite describing qualified immunity as “an affirmative defense” on a handful of occasions,218 the Court has also suggested that it represents the baseline presumption by describing the defense as “protecting . . . all but the plainly incompetent or those who knowingly violate the law.”219 The Court has also admonished trial court judges to be creative in using the Federal Rules of Civil Procedure to weed out claims before trial or extensive discovery to “protect[] the substance of the qualified immunity defense.”220

212 See id. at 201.
215 See id. at 319-20.
216 See id. at 309-10.
217 See id. at 316.
220 *Crawford-El v. Britton*, 523 U.S. 574, 597 (1998); see also id. at 597-601; *Harlow*, 457 U.S. at 819-20 n.35;
But what if the Court concludes that presumptions and stern talking-to’s are not enough to move trial courts towards greater recognition of qualified immunity? One option would be to find additional ways to expand the scope of appellate review. Although it is too soon to tell for certain, something along those lines may have happened last Term in \textit{Scott v. Harris},\footnote{127 S. Ct. 1769 (2007).} a case arising out of a high-speed police chase that rendered a man a quadriplegic.\footnote{See id. at 1173.} The trial judge had denied the defendant’s motion for summary judgment based on qualified immunity, but Justice Scalia’s majority opinion did not follow what even it described as the “usual[]” course of “adopting . . . the plaintiff’s version of the facts.”\footnote{Id.} Instead, it reasoned that because the record contained “a videotape capturing the events in question,” it was appropriate for the Justices to view the tape and draw their own conclusions about the true state of events.\footnote{See id. at 1775.} Although video records are obviously not available in all situations, \textit{Scott} is far from unique\footnote{See, e.g., Stewart v. Prince George’s County, 75 Fed. Appx. 198, 202-04 (4th Cir. 2003) (reversing district court’s denial of summary judgment in § 1983 case involving allegations of constitutionally excessive force based heavily on Court of Appeals’ review of a videotape of the encounter).} and an increasing number of law enforcement and other agencies appear to be videotaping encounters with the public.\footnote{See, e.g., Richard A. Jeo, \textit{The Impact of Miranda Revisited}, 86 J. CRIM. L. & CRIMINOLOGY 621, 682 (1996); Lucy S. McGough, \textit{Good Enough for Government Work: The Constitutional Duty to Preserve Forensic Interviews of Child Victims}, 65 LAW & CONTEMP. PROBS. 179, 186 (2002).} In addition, \textit{Scott}’s reasoning could, in principle, be applied to any situation in which a factual dispute is based in part on tangible evidence – including memos, correspondence, or emails – that reviewing courts would be capable of reviewing for themselves.

Finally, if the Justices conclude that expanded appellate monitoring will not accomplish their aims, one might see a majority overcome their apparent reluctance to impose more stringent limitations on trial courts’ initial decisions. Indeed, an effort to do so narrowly failed as recently as 1998. As noted above,\footnote{See supra text accompanying notes 207-11.} the Court’s decision in \textit{Harlow v. Fitzgerald} eliminated any consideration of a defendant’s subjective intent from the qualified immunity analysis. Yet there remain a few areas of constitutional law – most notably the Equal Protection Clause, “retaliation” claims under the First Amendment, and “cruel and unusual punishment” claims under the Eighth Amendment – where a plaintiff cannot establish an underlying constitutional violation without
proving that the defendant acted with an impermissible motive. The question in *Crawford-El v. Britton*, was how federal trial judges should deal with motions for summary judgment based on qualified immunity in those circumstances.

The answer was fairly clear under existing law. Consistent with orthodox summary judgment practice, the trial court should begin by resolving any factual disputes in favor of the plaintiff and constructing the most pro-plaintiff account reasonably supported by the record. Accordingly, so long as the plaintiff presents some evidence that can be read to suggest that the defendant acted with the requisite intent, the court should assume the intent's existence and then ask whether the defendant's actions violated the plaintiff's constitutional rights and whether that violation would have clear to a reasonable official as of the date in question.

But the problem with this approach, at least from the perspective of a Justice who thinks qualified immunity furthers important social policies, is that it can make it quite difficult to grant summary judgment. Because often “the required state of mind [will be] utterly inconsistent with a reasonable belief in the legality of one’s actions”—such as, for example, a defendant whom a court must assume *purposefully* discriminated against a member of a racial minority—resolution of the underlying factual issues in the plaintiff’s favor will often dictate denial of the defendant’s motion.

Although the *Crawford-El* majority acknowledged that this intersection of summary judgment procedures and the standards for granting qualified immunity posed certain problems, it also declined to do much about them. Reasoning that any further procedural changes would threaten to weed out meritorious suits and raise legitimacy concerns in the absence of action by Congress or the appropriate Rules Committee, Justice Stevens’ majority opinion expressed confidence that the best way to deal with the matter was to grant federal trial judges “broad discretion in the management of the factfinding process.”

---


233 See id. at 591

234 See id. at 594-97.

235 Id. at 601.
To the four dissenters, this “trust the district courts” approach was wholly unsatisfactory. In an opinion joined by Justice O’Connor, Chief Justice Rehnquist lambasted the majority for making a defendant’s ability to obtain a quick dismissal “dependant on the varying approaches of 700-odd district court judges” and proposed what might be best characterized as a broad rule of exclusion: Once a defendant offers a facially legitimate reason for the action being challenged, the Chief Justice argued, the only way a plaintiff should be able to avoid summary judgment is by producing “objective evidence[] that the offered reason is actually a pretext.” Justices Scalia and Thomas would have gone further and imposed a firm rule: Grant immunity whenever the defendant can offer any “objectively valid” reason for the challenged decision, regardless of any evidence of invidious intent.

Although these arguments did not carry the day, Crawford-El nicely illustrates the inescapable relationship between techniques for increasing Supreme Court control over lower courts and the underlying substantive goals a Justice hopes to accomplish. The five Justices in the majority may well believe that qualified immunity furthers important social purposes and that district courts are sometimes too slow to recognize the defense. But they also appear to believe, as Justice Kennedy stated in his concurring opinion, that “[p]risoner suits . . . can illustrate our legal order at its best,” and that dismissing a meritorious one at the pleading stage would represent a particularly bad outcome. In contrast, Justice Scalia acknowledged that his proposed rule would impose a “severe restriction upon ‘intent-based’ constitutional torts,” but wrote that he was “less put off by that consequence than some may be” because he believed that “no ‘intent-based’ constitutional torts would have been actionable under the § 1983 that Congress enacted.” Accordingly, future developments in this area may well be influenced as much by the current Court’s attitude towards the regime spawned by its 1961 decision in Monroe as its sense for how faithfully the lower courts are implementing the doctrine that has grown up around it.

236 Id. at 610 (Rehnquist, C.J., dissenting).
237 Id. at 602.
238 Id. at 612 (Scalia, J., dissenting).
239 Id. at 602.
240 Id.
C. Switching the Directionality: “Reasonable” Searches and Seizures

Despite their modern guise, the stories of the previous Sections were familiar in one respect: Like many of the most famous Warren Era developments, they involved the Supreme Court crafting doctrine for the purpose of broadening and protecting federal law defenses. But the link between distrust of lower courts and sympathy for federally protected rights is not inherent. In fact, as this Section explains, recent developments in an area closely associated with Warren Court activism now seem to be motivated, at least in part, by a desire to ensure that lower courts are not too generous to criminal defendants.

With the possible exception of desegregation litigation, it would be hard to think of an area in which struggles between the Supreme Court and lower courts have played as prominent of a role as the constitutional law of criminal procedure. During the Warren era, the Court not only constitutionalized (and thus federalized) huge swaths of criminal procedure, it also employed numerous techniques that that seemed designed to push trial courts to uphold more claims of federal rights and to facilitate meaningful appellate review when they failed to do so. The Court replaced a number of fuzzy standards with bright-line rules that favored defendants. It established strong presumptions that defendants wished to exercise their federal rights and that any violation of those rights was prejudicial and required reversal. It imposed rules of exclusion and enforced them via duties of explanation. And it both expanded its own ability to review state criminal convictions by loosening the “adequate and independent state law grounds” doctrine and enlisted the

---


243 See, e.g., Gideon v. Wainwright, 372 U.S. 335 (1963) (holding that every indigent criminal defendant charged with a felony is entitled to state-supplied attorney); Bruton v. United States, 391 U.S. 123 (1968) (holding that certain categories of accomplice confessions are per se inadmissible during a joint trial, thus overriding discretion that trial courts typically enjoy under Federal Rule of Civil Procedure 403 and comparable state rules).

244 See, e.g., Boykin v. Alabama, 395 U.S. 238 (1969) (overturning capital conviction because record did not affirmatively show defendant had knowingly and voluntarily waived several enumerated federal constitutional rights).

245 See, e.g., Chapman v. California, 386 U.S. 18 (1967) (holding federal constitutional errors require reversal unless shown to be harmless beyond a reasonable doubt).


lower federal courts’ aid by massively expanding the federal post-conviction writ of habeas corpus.248

Revolutions often generate backlashes,249 and this one was no exception. But although legal scholars have spent a great deal of time debating to what extent the Burger and Rehnquist Courts “retrenched” from the Warren legacy both in general250 and in criminal procedure in particular,251 I do not intend to join the fray. Instead, the remainder of this Section will focus on far more recent decisions in one particular area – doctrines governing the “reasonableness” of searches and seizures under the Fourth Amendment – and argue that they suggest an attempt to limit federal trial court’s ability to suppress evidence offered against criminal defendants.

One potentially serious barrier to controlling pro-defendant trial court rulings was removed long ago. Because the Fifth Amendment’s Double Jeopardy Clause252 has long been understood to bar government-taken appeals following an acquittal,253 strict enforcement of the final judgment rule in this context would have the effect of shielding many trial court rulings from any appellate scrutiny. In apparent response to this concern, both Congress and numerous state legislatures have enacted statutes authorizing interlocutory appeals from trial court rulings that grant a defendant’s motion to suppress evidence,254 and the Supreme Court has construed the federal statutes regulating its own jurisdiction as authorizing it to hear cases in which such interlocutory suppression rulings are affirmed on appeal.255

Merely authorizing appellate review, however, is not enough to ensure effective Supreme Court control over trial court decisionmaking. In addition,
the Court’s task is further complicated by the fact that even Justices who have expressed a strong general preference for formulating legal doctrines in terms of rules rather than standards have suggested that, given the near limitless variety of encounters between citizens and police, a number of basic Fourth Amendment concepts are “not readily, or even usefully, reduced to a neat set of legal rules.”

Based on statements like this, it would be easy to surmise that Fourth Amendment jurisprudence is some sort of rule-free zone. But that would be a serious mistake. In fact, during the last decade alone, the Supreme Court has announced and expanded numerous clear rules in the Fourth Amendment context. It is “reasonable” per se, the Court held in 1999, to pull over a motorist whenever an officer has probable cause to believe that she has violated any traffic regulation, regardless of circumstances or severity. Once a car is pulled over, the Court has further held that officers may always order both the driver and any passengers out of the vehicle. In 2001, a five-Judge majority held that the Fourth Amendment permits arresting any driver whom an officer has probable cause to believe committed any crime in the officer’s presence – including misdemeanor offenses such as failure to wear a seatbelt. And once an officer has decided to arrest, the Court has held that she enjoys an absolute entitlement to search the car’s entire passenger compartment as an “incident” to that arrest—a holding that the Court in 2004 extended to cover a situation in which the suspect was handcuffed in the back of a police cruiser at the time of the search.

Given the sheer number of rules of the road and the ubiquity with which we all violate them, these rulings render it constitutionally permissible for police officers to stop and search the passenger compartment of almost any

256 See generally Scalia, supra note 50.

257 United States v. Sokolow, 490 U.S. 1, 8 (1989) (discussing concepts of “reasonable suspicion” and “probable cause”) (internal quotation marks and citation omitted); see also Scott v. Harris, 127 S. Ct. 1769, 1777-78 (2007) (discussing concept of “reasonable force”); Scalia, supra note 50, at 1186.


261 See Maryland v. Wilson, 519 U.S. 408 (1997).


car at almost any time. Of course, that fact, standing alone, does not prove that these rulings were motivated by a desire to control lower courts. Indeed, it might seem equally plausible that their aim was the one invariably cited in the Court’s opinions: the need to provide clear guidance for police officers in the field.

Although this account has a great deal of force, there is a serious problem with treating it as the sole impetus for the Court’s recent decisions. Cops no doubt benefit from being given a list of things that they may do. But if – as the Court has repeatedly stated – the basic idea behind the exclusionary rule is to use the threat of suppression to deter undesirable conduct, there are also no doubt circumstances in which the police would benefit from being given a list of things that they may not do. But despite this fact, the Supreme Court has been steadfast in refusing to create – or permitting lower courts to begin to develop – any such lists. Indeed, what is perhaps most notable about the Court’s recent use of rules in the Fourth Amendment context is their one-way character: the great bulk of them hold that a particular type of police conduct is per se acceptable under the Federal Constitution. More than perhaps any characteristic, this suggests that what may also be going on is an attempt to make it more difficult for lower courts to suppress evidence based on Fourth Amendment violations.

But what about situations for which a majority of Justices have found themselves unable to agree on a rule of per se compliance with the Fourth Amendment? There, recent decisions show use of two different techniques for controlling lower court decisionmaking that mirror those discussed in the previous Sections, as well as a third that illustrates a far more direct method for preventing lower courts from granting relief based on disfavored rights.


266 See, e.g., Thornton, 541 U.S. at 622-23; Atwater, 532 U.S. at 347; Belton, 453 U.S. at 458.


269 The main exceptions are two recent decisions holding that certain police actions always constitute a “search” or “seizure.” See Brendlin v. California, 127 S. Ct. 2400 (2007) (passengers in a car are “seized” at the time the driver pulls over in response to a show of police authority); Kyllo v. United States, 533 U.S. 27 (2001) (use of imaging technology to gather details about interior of a home not perceivable to an ordinary observer is a “search”). In each of these circumstances, however, the conclusion dictated by the rule does not compel exclusion of the resulting evidence, it simply directs the trial court to conduct additional Fourth Amendment analysis.
For one thing, the Court has both emphatically stated and steadily expanded one very important rule of exclusion. As I explained earlier, a trial court’s ability to find and characterize the underlying historical facts may pose the single greatest hindrance to effective appellate monitoring, and this problem is particularly acute when the underlying legal standard invites assessments about someone’s state of mind. Accordingly, it should come as little surprise that just as the Court eliminated such considerations from qualified immunity analysis, it has repeatedly stated that a police officer’s beliefs, motives, or intentions generally “play no role . . . in Fourth Amendment analysis.”

The Court also directed intensified appellate monitoring in its critically important, but little noted, decision in *United States v. Ornelas.* Although the Court has repeatedly emphasized the highly fact-dependent and case-specific nature of both sorts of determinations, *Ornelas* nonetheless held that a trial court’s assessment of the legal significance of any particular fact and its conclusions about whether the facts as a whole establish reasonable suspicion or probable cause should be reviewed *de novo* on appeal. Any other standard of review, Chief Justice Rehnquist wrote, risked creating situations where otherwise similar cases could be resolved differently depending how “different trial judges draw general conclusions.” And although it acknowledged that even *de novo* review had only a limited ability to “unify precedent” in the absence of clear rules, the Court stressed – in language upon which it would later rely to justify the same conclusion in the punitive damages context – that “[i]ndependent” appellate review was “necessary if appellate courts are to maintain control of, and to clarify, the legal principles.”

Even more recently, the Supreme Court has reminded us that there is another technique it can employ if it concludes that lower courts are too quick to grant relief based on a claimed violation of federal rights. In 2006’s *Hudson*

---

270 See supra note 116 and accompanying text.

271 See supra notes 207-10 and accompanying text.


274 See supra note 268 and accompanying text.

275 See 517 U.S. at 699.

276 Id. at 697 (quoting *Brinegar v. United States*, 338 U.S. 160, 171 (1949)).

277 See supra note 152 and accompanying text.

278 See *Ornelas*, 517 U.S. at 697 (emphasis added).
a five-Justice majority held that violations of the Fourth Amendment’s “knock-and-announce” rule do not require suppression of any evidence seized during the subsequent search. Although it went uncited in Justice Scalia’s majority opinion, Hudson bears a close resemblance to the Court’s decision thirty years earlier in Stone v. Powell. Ever since Brown v. Allen, the general rule has been that prisoners who believe their federal constitutional rights were violated during state court criminal proceedings may seek relief from the lower federal courts. In Stone, however, the Court created a special rule for Fourth Amendment claims, holding that arguments that unconstitutionally seized evidence was wrongly admitted at the defendant’s trial are simply not cognizable in federal habeas proceedings. Stone’s approach has not spread to other types of claims in the habeas context, and only time will tell if Hudson has longer legs. Yet both cases demonstrate the basic point that the ultimate way of preventing too frequent vindication of a disfavored right is to bar the granting of a remedy.

Especially in conjunction, these various developments suggest a concerted effort by the Supreme Court to discourage lower courts from suppressing evidence based on Fourth Amendment violations. At the same time, other aspects of the Court’s recent behavior underscore the challenges the Justices face in seeking to do so. Although the Court is now regularly deciding just 80 cases per year, during its last three Terms alone it has issued thirteen opinions regarding the constitutional rules governing “searches and seizures,” all but one of which it has decided in favor of the government litigant. Because the Court reverses or vacates far more often than it affirms, these numbers suggest that the Justices still perceive the need to spend a relatively large

281 344 U.S. 443 (1953).
282 See id. at 494.
285 See supra note 33.
286 See 2005 Statistics, supra note 33, at 382, tbl. III (listing four “search and seizure” cases, three of which were decided in favor of government); 2004 Statistics, supra note 33, at 429 tbl. III (three cases; all decided in favor of government); 2003 Statistics, supra note 363, at 507, tbl. III (six cases; all decided in favor of government).
287 See 2005 Statistics, supra note 33, at 382, tbl. II(D) (stating Court reversed or vacated lower court’s judgment in 58 of 81, or 71.6% of, cases it reviewed on a writ of certiorari); 2004 Statistics, supra note 33, at 426, tbl. II(D) (60 of 83, or 72.2%); 2003 Statistics, supra note 33, at 507, tbl. II(D) (59 of 76, or 77.6%)
amount of their own time engaged in fact-bound review of lower court
decisions that resolve Fourth Amendment issues in favor of criminal
defendants, perhaps in order to provide greater “analogical anchoring” for
future cases.\(^{288}\)

\* \* \*

The subject areas discussed in the previous three Sections could easily be
expanded, but it is not my purpose to catalogue all of the spheres in which an
attempt to shape and direct lower court decisionmaking is a plausible
explanation for recent developments in Supreme Court doctrine. Instead, my
aim in this Part has been to make three fairly straightforward points:

\textbf{First}, distrust of lower courts as a motivator for doctrine creation did not
end with the Warren Court and need not be spurred by a desire to enact
policies likely to please the political left. To the contrary, one obvious
connection between all three subjects discussed in this Part is that the Court’s
efforts have been quite consistent with the tenets of political conservatism.

\textbf{Second}, although use of rules may be the technique most commonly
associated with efforts to control front-line decisionmakers, it is far from the
only tool at the Supreme Court’s disposal. Rules were not the sole tactic
employed in any of the areas just discussed, and it would be a stretch to label
use of rules the primary strategy in any of them.

\textbf{Third}, the various techniques discussed in Part II are in large measure
complementary to – and, to a certain extent, substitutes for – one another. In
all of the areas just discussed, the Court has employed multiple techniques, and
at times is has acknowledged that it is using one because another is not
available.\(^{289}\)

\textbf{CONCLUSION}

Legal scholars who study the relationship between rights and remedies are
fond of saying that the real meaning and value of the former are determined by
the generosity or stinginess of the latter.\(^{290}\) This Article suggests an important
corollary: The extent to which the Supreme Court values the rights it espouses
in its opinions is demonstrated in large measure by the lengths to which it is
willing to go to secure lower court compliance with them. The fact that the

\(^{288}\) See Carolyn Shapiro, \textit{The Limits of the Olympian Court: Common Law Judging versus Error
deciding a series of cases in which it applied a standard, the Court could harness the traditional
common law method of analogical reasoning to mark a path for lower courts”).

\(^{289}\) See, e.g., supra text accompanying notes 274-78 (discussing Court’s adoption of \textit{de novo}
review in the Fourth Amendment context).

\(^{290}\) See, e.g., Levinson, supra note 284, at 858 (“Rights are dependent on remedies not just
for their application to the real world, but for their scope, shape, and very existence.”)
Justices of the Warren era were willing to bear enormous costs to achieve even partial compliance with their vision of a better, fairer criminal justice system suggests that they were enormously committed to that goal. In contrast, various rulings by the Burger, Rehnquist, and now Roberts Courts that make it more difficult to overturn criminal convictions based on even conceded constitutional violations indicate a comparatively lesser commitment.291

This Article also holds important lessons for political scientists. Its exploration of the various ways in which the Supreme Court can craft legal doctrines to constrain and direct the course of lower court decisionmaking may suggest a partial answer to why every study of which I am aware has found high levels of lower court compliance.292 More broadly, its demonstration that the power dynamic between trial and appellate courts is far from static illustrates the critical need “to take law and legal institutions seriously.”293

In addition, my analysis suggests at least two areas for future research. First, this Article’s largely descriptive account raises important normative questions both about the relationship between the Supreme Court and lower courts, as well as appropriate methodologies in constitutional and statutory interpretation. A number of important pieces have examined the issue from the lower courts’ perspective, in particular, the source, extent, and nature of their “obligation” to follow Supreme Court precedent.294 But there has been very little work addressing what may be seen as the opposite question: the extent to which it is legitimate for Supreme Court Justices to interpret concededly binding legal materials with at least one eye focused on controlling lower court decisionmaking in cases that have not yet arisen and will likely never come before the Court itself.295

291 I am grateful to Brandon Garrett for first bringing this point to my attention. See also Steiker, supra note 251, at 2468-70.
292 Citations to a number of the studies are collected in Kim, supra note 3, at 395-96 & nn. 41-47, and Songer, supra note 3, at 43-46.
293 Friedman, supra note 19, at 262.
295 This issue has a certain overlap with, but is nonetheless distinct from, debates about the existence and scope of the Supreme Court’s “supervisory power.” See, e.g., Barrett, supra note 73. Questions about the supervisory power generally involve whether and to what extent structural features of the Constitution and aspects of our political history authorize the Supreme Court to impose restrictions on lower federal courts without claiming that those restrictions are required by some external source of authority, such as the Constitution or a
Second, although this Article has focused almost exclusively on doctrine formulation by the Supreme Court in its adjudicatory capacity, another profitable area for further inquiry would be the degree to which concerns about lower courts may shape and be visible in the products of both rule-making and legislation. As to the former, consider the 1994 amendments to the Federal Rules of Evidence that carved out an exception to trial judges’ ability to exclude any evidence that they deem unduly “prejudicial” by declaring evidence about certain categories of criminal convictions *per se* admissible\(^\text{296}\) and others to the Federal Rules of Civil Procedure that created a new exception to the final judgment rule for district court decisions about whether to certify a class action.\(^\text{297}\) As to the latter, consider the provisions of the Antiterrorism and Effective Death Penalty Act that declare that only decisions by the Supreme Court itself may be relied upon as a basis for granting federal habeas relief,\(^\text{298}\) and make federal district courts’ ability to entertain second-or-successive habeas petitions contingent on approval by the Courts of Appeal,\(^\text{299}\) the sections of the Prison Litigation Reform Act that require lower courts (but not the Supreme Court) to act on petitions to modify injunctive decrees within a particular time,\(^\text{300}\) the parts of the Class Action Fairness Act that dramatically expand the scope of state law actions that may be removed to federal court,\(^\text{301}\) or pretty much the entire history of the Sentencing Reform Act\(^\text{302}\) and the Federal Sentencing Guidelines.\(^\text{303}\) Worries about lower courts, these examples suggest, are hardly confined to the Supreme Court of the United States.

\(^{296}\) See *Fed. R. Evid.* 413 (enacted by Congress in 1994).


\(^{299}\) *Id.* § 2244(b)(3).

\(^{300}\) 18 U.S.C. § 3626(c)(2).

\(^{301}\) 28 U.S.C. §§ 1332(d), 1453(b) & (c).

\(^{302}\) 18 U.S.C. § 3551 *et seq.*