ALAN B. MORRISON*

The Necessity of Tradeoffs in a Properly Functioning Civil Procedure System

INTRODUCTION

When judges, lawyers, and law professors discuss tradeoffs, it is usually in the context of debates about substantive policies.

* Lerner Family Associate Dean for Public Interest & Public Service, George Washington University Law School.

Will too much environmental regulation make our industries uncompetitive? Will restricting the grounds for which an employee may be fired limit the flexibility of management to control the workplace? Will increasing tort liability for dangerous products stifle innovation? Do certain provisions in the tax code unduly favor one industry over another? In each case there are substantive policies that would be advanced or hindered by taking one position or the other. Sometimes the courts or legislatures are explicit about the tradeoffs; at other times they are not.

But procedure seems different, at least at first blush, perhaps in part because in the federal system the rules are issued by the Supreme Court after a lengthy committee process involving judges and lawyers, rather than elected legislators. How can the form of a complaint or the time to answer or amend involve tradeoffs in any meaningful sense of that word? Discovery rules can be viewed as simply the means by which information is obtained for use at trial, and if there are tradeoffs, they are not apparent on the face of the rules. That impression may explain why procedural rules seem so bland, and why they are so hard to understand unless the tradeoffs are made visible and their bases, along with the reasons why one choice rather than another was made, revealed.

Rule 1 of the *Federal Rules of Civil Procedure* is a perfect illustration of hidden tradeoffs. It directs the courts to do what every litigant, lawyer, and judge would support: to administer the Rules "to secure the just, speedy, and inexpensive determination of every action and proceeding."[1] The problem is that just results often come slowly or expensively. Or, conversely, a speedy result may not be a just one, and even inexpensive cases are not always speedy. The good news is that courts and parties rarely rely on Rule 1, and when they do, it is generally as window dressing to support a result reached under another Rule. But to be accurate, Rule 1 should be recast to require the courts to provide a "just determination of every action" and to do so with "appropriate speed and without undue expense" under the circumstances. Doing that would bring it in line with one of the relatively few Rules where the tradeoff is explicit, Rule 26(b)(2)(C)(iii). Under that Rule, the court is directed to limit discovery if "the burden or expense of the proposed discovery outweighs its likely benefit, considering the needs of the case, the amount in controversy, the parties' resources, the importance of the

[1] FED. R. CIV. P. 1.

issues at stake in the action, and the importance of the discovery in resolving the issues."[2] Whether that actually helps judges decide real cases is another question, but the Rule surely frames the tradeoffs that the drafters considered appropriate in resolving discovery disputes.

I try to show my Civil Procedure students that most rules are written to achieve some purpose or to solve some problem that arises in litigation. I then advise them that, unless that purpose or problem can be divined, the meaning and operation of the rule and any exceptions to it cannot be understood. Unfortunately, one purpose can rarely be advanced without some other purpose being set back, or reduced in significance. That means that there must be a tradeoff, hopefully consciously and openly made, even if the evidence of the tradeoff is not apparent to most observers or is generally not something on which lawyers or judges focus in using a rule of procedure. Yet, for the law student trying to grasp the significance of a rule, looking for the tradeoff and appreciating why it was made are the surest ways to master a rule and learn how it should be applied. Put another way, even the most vanilla-sounding rules are not "neutral" because they generally help one side more than the other, even if that is not apparent from the face of the rule. To be sure, some tradeoffs are harder to locate than others, and some rules involve a tradeoff in only the most theoretical application of that term. But, by and large, the search for a tradeoff is far more likely to be a fruitful tool for the student of Civil Procedure than is the assumption that a rule of procedure serves no more purpose than does the rule that the pitcher's mound in baseball shall be exactly sixty feet, six inches from home plate.

The common law in fields such as torts, contracts, and property was developed on a case-by-case basis, which meant that the substantive law was often determined by the facts (which may be more favorable for one side than the other). A major downside of the common law is that the outcome is never certain, making compliance and planning more difficult for all. For most procedural rules, the value of at least a reasonable degree of certainty is often seen as an overriding consideration on the theory that generally a party can comply with whatever rule there is, so long as it is known in advance. That explains why the procedures by which cases are handled are found in rules or statutes rather than developed on a case-by-case

---

[2] *Id.* 26(b)(2)(C)(iii).

basis, as is the common law, although the presence of rules does not eliminate disputes over their meaning.

Before illustrating some of the most significant tradeoffs in the *Federal Rules of Civil Procedure*, a few other points are worth noting because they apply to a number of the specific rules that will be discussed. First, the Rules are supposed to be trans-substantive, which is a fancy way of saying that they are supposed to be applicable to all the different types of substantive-law claims that are litigated in the federal courts. There are some exceptions in the Rules themselves, such as the requirement for greater specificity in pleading fraud or mistake in Rule 9(b), and Congress has introduced a heightened pleading requirement in complaints alleging violations of the federal securities acts. In some respects the one-size-fits-all approach seems odd given the very different substantive policies involved and the substantive tradeoffs made in different substantive areas of law. But the goal of having trans-substantive rules can be defended as itself a form of tradeoff: it is simpler to have a single set of procedural rules for all areas of the law and, therefore, the label attached to a cause of action will not have great significance, even if the rules work better for some types of claims than others.

Second, the tradeoffs in the Rules are not fixed, but have been re-calibrated as circumstances change. As discussed below, discovery may be the clearest example of how the Rules have evolved as discovery has become much more significant over the decades. Most recently, electronic record keeping made it possible to discover the previously undiscoverable—albeit at considerable burdens of time and expense—thereby suggesting a need for a different balancing among the competing interests in discovery.

One of the most dramatic examples of the changing nature of the tradeoffs made in the Rules is found in Rule 26(a), which imposes on each party the duty to make certain affirmative disclosures. This obligation, first instituted in 1993 on an optional basis for each district, was added in an effort to reduce costs and lessen delays, and in doing so altered the adversary system so that parties became obligated to do more than simply respond to requests made by the other side. Those obligations were lessened in 2000 and made uniform for all district courts, as the Rules Committee sought to find the proper balance. Even with those changes, the new Rule represents a significantly different tradeoff than did the original version.

Third, there are tradeoffs in the type of procedural rule that is chosen between those that create bright lines and those that instruct

the judge to decide the question based on the specific facts of the case before her. Consider two alternative approaches that different rules involving time actually embrace. A defendant is given a specific number of days to answer the complaint or to respond to a motion for summary judgment, whereas a plaintiff may amend her complaint after the initial grace period in Rule 15(a)(2) "when justice so requires."[3] Similarly, Rule 15(c)(1)(C)(i) allows relation back of an amended complaint to add a defendant in certain circumstances, provided that the defendant "will not be prejudiced" thereby. And a motion to intervene under Rule 24 will be granted if it is "timely," which has been held to include intervention even after a final judgment has been entered.[4]

The type of rule chosen itself contains a tradeoff between greater certainty and greater fairness, which some might call greater flexibility. And, while lawyers are capable of finding grounds to litigate the meaning of even those Rules in which the time is set in a precise number of days, the decision to focus on prejudice or timeliness is almost certain to generate more litigation than one that provides a fixed number of days within which some action must be taken, but with a greater likelihood of achieving a just result in a particular situation.[5]

This Article focuses mainly on tradeoffs contained in the *Federal Rules of Civil Procedure*, which are issued by the U.S. Supreme Court, with most of the work done by committees of judges, practicing lawyers, and law professors as part of a very public and open process. It also discusses statutes enacted by Congress, mainly as they affect the jurisdiction of the courts. Those statutes also have significant impacts on the outcome of disputes and, not surprisingly, contain tradeoffs as well. And finally, Article III of the Constitution, which creates limited jurisdiction for the federal courts, is the most fundamental tradeoff because it denies the vast majority of lawsuits a federal forum and instead prefers state courts as the basic locus for litigation.

---

[3] *Id.* 15(a)(2).

[4] *Id.* 24(a); Smuck v. Hobson, 408 F.2d 175, 182 (D.C. Cir. 1969).

[5] Similar choices apply in substantive areas where Congress has chosen to make the tax laws very specific and the antitrust laws much more general.

I

PRE-SUIT

A. Subject Matter Jurisdiction—Access to the Federal Courts

Article III of the Constitution allows federal courts to decide cases or controversies only in a limited set of circumstances—itself a tradeoff—of which the most important are those cases "arising under this Constitution, the Laws of the United States, and Treaties made . . . under their Authority" (federal question jurisdiction) and controversies "between Citizens of different States" (diversity of citizenship jurisdiction).[6] Article III also provides that Congress shall determine the jurisdiction of the lower federal courts; Congress has chosen not to grant those courts the full extent of the power available under Article III.

When Congress created the federal trial courts under the first Judiciary Act, it chose not to allow them to decide federal question cases, but limited their role to diversity cases. Eventually, in 1878, Congress decided that the added burden on the federal courts of dealing with federal question cases was more than offset by the benefit of allowing federal, rather than only state courts, to make the initial decisions in cases arising under federal law in the first instance.[7] But Congress did not open the federal courts to all such cases, choosing instead to limit jurisdiction to cases where the amount in controversy exceeded a certain sum, initially $500, which was raised first to $2000, then to $3000, and finally to $10,000 in 1958.[8] Then in two steps, first in 1976 and then in 1980, Congress reversed itself and eliminated the amount in controversy requirement in federal question cases.[9] In effect, it concluded that keeping some federal question cases out of federal court was a bad tradeoff, because it meant that state courts were deciding questions of federal law (subject to possible review in the Supreme Court) and because the burden on the courts of determining whether some federal constitutional claims

---

6 U.S. CONST. art. III, § 2, cl. 1.

7 Act of Mar. 3, 1875, ch. 137, § 1, 18 Stat. 470.

8 Act of July 25, 1958, Pub. L. No. 85-554, § 4, 72 Stat. 415.

9 Act of Oct. 21, 1976, Pub. L. No. 94-574, § 2, 90 Stat. 2721; Act of Dec. 1, 1980, Pub. L. No. 96–486, § 2, 94 Stat. 2369.

were "worth" more than $10,000 was not worth the benefit of keeping a few "small" cases out of the federal courts.[10]

For diversity cases, Congress has continued its gatekeeper role by limiting federal court jurisdiction to larger cases, currently those in excess of $75,000.[11] The generally accepted reason for diversity jurisdiction in the Constitution is fear that state courts would unduly favor local citizens, whereas federal courts, in part because of the method of selecting federal judges, would be less inclined to be biased in favor of local citizens.[12] In theory, that would suggest that all diversity cases can be heard in federal court, but Congress has consistently rejected that conclusion, largely because, given the limited number of federal district judges, they would be overwhelmed if every diversity case could be brought in federal court. Presumably, local bias would also be present even in cases involving less than $75,000, but Congress decided, at least implicitly, that the tradeoff of the burdens on the system of allowing all such cases to be heard in federal court was not worth the gain in neutrality to the out-of-state litigant in smaller cases.

Another way that Congress has limited diversity jurisdiction is through the rule of complete diversity. Under complete diversity, all parties on one side must be citizens of states different from all parties on the other side. The statutes governing diversity jurisdiction do not mention complete diversity, but the Supreme Court interpreted them to require it.[13] Congress has never rejected that reading, but it has created exceptions where it found that the general tradeoff gained from limiting federal courts to complete diversity was not justified. A prominent example of Congress reaching a different balance on federal versus state court for diversity cases is found in the Federal Interpleader Act of 1917,[14] under which the federal courts have

---

10 Fifth Ave. Peace Parade Comm. v. Hoover, 327 F. Supp. 238, 242 (S.D.N.Y. 1971) ("Certainly they may be difficult of evaluation, but 'priceless' does not necessarily mean 'worthless.'").

11 28 U.S.C. § 1332(a) (2006).

12 A similar rationale also explains why the Constitution gives aliens access to federal courts, but it does not explain why Congress allows U.S. citizens to choose a federal forum to sue an alien. The same mismatch occurs in diversity cases where a citizen can bring a case in federal court in her home state when there is no arguable prejudice to the plaintiff if the case were heard in her own state court, but a defendant cannot remove a case to federal court if it is brought in a state court in his home state. *See* 28 U.S.C. § 1441(b).

13 Strawbridge v. Curtiss, 7 U.S. (3 Cranch) 267, 267–68 (1806).

14 Federal Interpleader Act of 1917, ch. 113, 39 Stat. 929 (current version at 28 U.S.C. § 1335).

jurisdiction so long as there is diversity between any two adverse claimants. The Act was passed to deal with situations in which there are several claimants to a fixed fund, often an insurance policy, and the claimants live in different states, such that it would be difficult, if not impossible, to obtain personal jurisdiction over all of them in any state court. Because the claims do not arise under federal law, the only way to get the cases into a federal forum is through diversity, and even that might be thwarted if complete diversity were required and two of the claimants were citizens of the same state. Because Congress concluded that the benefits of having a federal forum to resolve these interstate matters was more important than limiting access to the federal courts by mandating complete diversity, it allowed minimal diversity and set the amount in controversy at $500.[15] In short, Congress simply made a different tradeoff for that limited set of cases.

In the late 1980s, Congress was faced with a similar kind of choice about whether to relax some of the rules on subject matter jurisdiction in cases involving multiple parties where there was subject matter jurisdiction as to some but not all of the parties or claims. The U.S. Supreme Court had decided a series of cases, generally finding that these additional parties could not be added,[16] although the Court was more generous with additional claims against existing parties.[17] The result was that some parties were forced to litigate the same basic dispute in both federal and state courts,[18] which produced additional costs and some procedural unfairness. It was undisputed that, because there was subject matter jurisdiction for at least one claim in each case, there was no constitutional barrier to relaxing the rules on adding parties or claims; rather, the issue was which tradeoffs were appropriate to make, recognizing that any relaxation of these requirements would add to the workload of the federal courts by some probably indeterminable amount.

Congress took two different routes, depending on the basis for subject matter jurisdiction. If there is federal question jurisdiction, then parties and claims could be added if they were part of the same case or controversy.[19] But for diversity cases, the next subsection specified certain rules under which claims and parties could not be

---

15 *Id.*

16 *See, e.g.*, Aldinger v. Howard, 427 U.S. 1 (1976).

17 *See* United Mine Workers v. Gibbs, 383 U.S. 715 (1966).

18 *See* Finley v. United States, 490 U.S. 545 (1989).

19 28 U.S.C. § 1367(a).

added in diversity cases, thereby maintaining the old balance, which meant that efficiency of litigation gave way to keeping cases (or at least some parties and some claims) out of the federal courts.[20] Unfortunately, Congress did a very poor job of drafting and left out of the list of Federal Rules for which joinder was barred the class action rule—Rule 23—and some, but not all applications of Rule 20. Cases involving those provisions went to the Supreme Court, which held that the failure to include those provisions in § 1367(b) meant that the general rule applicable to non-diversity cases applied, and so additional, non-qualifying parties could be joined under those rules.[21] As the dissent pointed out, viewed from the perspective of congressional purposes, it is hard to see why Congress would have wanted to bring into the federal courts the much larger group of more complex class actions but exclude much smaller cases affected by § 1367(b), which is what the majority said Congress had done.[22] Put another way, the basic tradeoff in subsection (b) of less efficiency for diversity cases is understandable as a general proposition but seems very difficult to defend while allowing some, but not all, much more burdensome cases, such as class actions, to come to federal court.

The class action ruling under the supplemental jurisdiction statute will have only minor practical consequences because, while the Court was considering *Exxon Mobil Corp. v. Allapattah Services, Inc.*, Congress passed the Class Action Fairness Act (CAFA)[23] that greatly expanded the ability of parties to litigate state-law-based class actions in federal court. Prior Supreme Court rulings interpreted the diversity jurisdiction provision to require that all members of the class have claims of more than $75,000,[24] which meant that very few class actions based on state law could be heard in federal court. In the 1970s, it was generally plaintiffs who wanted to expand class action jurisdiction in federal courts, because they saw them as more favorable fora, but by the twenty-first century, it was the defendants who were seeking refuge in federal courts from class actions filed in state courts, often in venues that were considered very pro-plaintiff. Defendants generally thought that federal courts were less likely to

---

[20] *Id.* § 1367(b).

[21] Exxon Mobil Corp. v. Allapattah Servs., Inc., 545 U.S. 546 (2005).

[22] *Id.* at 593 (Ginsburg, J., dissenting).

[23] Class Action Fairness Act of 2005, Pub. L. 109-2, 119 Stat. 4 (codified at 28 U.S.C. § 1332(d)).

[24] *See, e.g.*, Zahn v. Int'l Paper Co., 414 U.S. 291, 301 (1973).

grant class certification than were at least some state courts (mainly those where plaintiffs elected to file). Defendants also wanted to be able to consolidate cases where there were multiple similar actions filed in different states, which was possible only if the cases could be removed to the federal courts. One barrier to consolidation was that the federal removal statute applicable to diversity cases applied only where none of the defendants was a citizen of the state in which the case was filed.[25] That limitation was itself a tradeoff between allowing a diversity defendant the choice of removal and precluding it when the defendant was in its home state, where presumably it would not be subject to local bias.

CAFA produced a major change in the diversity tradeoff.[26] First, as to amount in controversy, it became $5 million total (instead of $75,000 for each class member), provided that there were one hundred or more class members.[27] Second, for diversity, CAFA opted for minimal diversity, meaning that if one plaintiff class member and one defendant are diverse, the citizenship of the remainder of the class and of the other defendants is irrelevant. And third, one defendant can remove without requiring all defendants to join in the removal.[28] Thus, for a variety of reasons, including, for some supporters, the desire to limit the effectiveness of class actions, Congress decided that the benefits of making available a federal forum for major class actions offset concerns about flooding the federal courts with state-law cases. Congress attempted to soften the effect of the change by providing for remands in some circumstances;[29] however, the extent of the ability of plaintiffs to obtain remands is seriously in doubt, and in all likelihood, those provisions will prove to be relatively unimportant.[30]

A final tradeoff in this area involves the question of how to define the "citizenship" of a corporation. Congress first made the obvious choice of the state of incorporation and then added principal place of business.[31] Because of the rule of complete diversity, adding

---

[25] 28 U.S.C. § 1441(b).

[26] *See id.* § 1332(d).

[27] *Id.*

[28] *Id.* § 1453.

[29] *Id.* § 1332(d)(2), (3).

[30] A slightly different set of tradeoffs involving mass accidents is contained in 28 U.S.C. § 1369. That provision relaxes the rule of complete diversity where there are at least seventy-five natural persons who died (not merely injured, although injured persons may intervene) in a single discrete accident in certain prescribed situations.

[31] *Id.* § 1332(c)(1).

principal place of business reduced the opportunities for corporations to sue or be sued in federal court. Moreover, and perhaps more significantly, it precluded them from removing cases to federal court if they were sued where they had their principal place of business. The term principal place of business could have several meanings, and in *Hertz Corp. v. Friend*, the Court decided that the more readily ascertainable "nerve center" test was the proper one, in part because it concluded that Congress wanted as close to a bright-line rule as possible in this area in order to minimize preliminary and collateral litigation.[32] The choice was in some respects between certainty and simplicity on one hand, and complexity and, in some cases, greater fairness and adherence to the concerns about prejudice against out-of-state persons, which animates the diversity rationale, on the other, with tradeoffs present under both options.

B. Venue and Related Issues

Once a plaintiff has decided whether to file in federal or state court, and assuming that the defendant is amenable to suit in a variety of locations, 28 U.S.C. § 1391 imposes some modest restrictions on where the case may be brought.[33] Congress has also concluded that, even though venue is supposed to be a proxy for convenience—notice the similarity of roots of the two words—there are some cases in which the forum chosen meets the venue and personal jurisdiction requirements, but the case should nonetheless be transferred to another forum where the case might have been brought "[f]or the convenience of parties and witnesses, in the interest of justice."[34] The inquiry is very fact-dependent, and the burden is on the party seeking the transfer to overcome the presumption that the plaintiff, who has the burden of proof, should be deprived of his chosen forum. Despite these obstacles, some cases can be transferred to another federal court because Congress has made the judgment that, in those situations, the desire of the plaintiff to litigate in one place must take a backseat to the convenience of other parties and witnesses.

Another tradeoff in this area was created for quite different reasons and is limited to pretrial proceedings. In 28 U.S.C. § 1407, Congress created the Judicial Panel on Multidistrict Litigation, which has the

---

32 Hertz Corp. v. Friend, 130 S. Ct. 1181, 1193–94 (2010).

33 28 U.S.C. § 1391.

34 *Id.* § 1404(a).

power to transfer cases involving common questions of fact to a single court. The rationale for this provision is to assure consistency and increase efficiency by consolidating pretrial matters—ranging from discovery to class certification to motions for summary judgment—before one federal judge. The interests of individual plaintiffs (and perhaps even some defendants) in handling these pretrial matters in a more convenient forum are submerged to the efficiency interests of the courts and all the parties with similar claims in centralized litigation. In *Lexecon Inc. v. Milberg Weiss Bershad Hynes & Lerach*, the Court ruled that § 1407 required a retransfer back to the original forum for trial (perhaps because it was thought to be fairer to the plaintiff to return to her chosen forum).[35] While Congress has debated changing that rule, to date it has not acted, meaning that the district judge who is most familiar with the case cannot try it.

II
BEGINNING THE LAWSUIT

A. The Complaint

The drafters provided in Rule 8(a)(2) that a federal court complaint need contain only "a short and plain statement of the claim showing that the pleader is entitled to relief,"[36] a practice commonly referred to as "notice pleading." The Rule also requires that the complaint include a statement as to the basis for subject matter jurisdiction, but because federal courts are courts of limited jurisdiction, that is hardly a surprising or burdensome requirement. Indeed, including it serves as a check to remind counsel that they must satisfy that requirement or the case will be dismissed. The complaint need not include anything on personal jurisdiction or venue—which are also bases for dismissal or transfer—perhaps because they can be waived, unlike subject matter jurisdiction. The sample forms also show how simple a complaint can be, although many lawyers choose to make them more complex (and for some lawyers *much* more complex) than is needed. The relief sought does not have to request a specific dollar amount, although if the basis of jurisdiction is diversity of citizenship, it must at least allege that the requisite amount in controversy is met.

---

35 Lexecon Inc. v. Milberg Weiss Bershad Hynes & Lerach, 523 U.S. 26 (1998).

36 FED. R. CIV. P. 8(a)(2).

And even though a complaint initiates a *law*suit, it need not contain citations to the applicable law, although it often does.

Beyond permitting complaints to be simple, there are two related tradeoffs that have an important role in enabling plaintiffs to start a lawsuit in federal court.[37] First, at least in many cases, the complaint can be very general in its allegations regarding, for example, the cause of the injury to the plaintiff. It is enough to assert that the defendant drove his car negligently in striking the plaintiff's vehicle, without specifying that the defendant drove too fast, ran a red light, or was talking on his cell phone. Even if the complaint is specific, it can be amended to add new specifics or delete others as the case proceeds. Similarly, and as a practical matter of greater importance, especially in a case where the defendant has exclusive access to most of the important evidence on what caused the plaintiff's injury, a general allegation of, for example, defective design or manufacture of a product, will suffice. This latter example illustrates a tradeoff in the Rules that makes it easier—and in some cases actually makes it possible—for a plaintiff to start his lawsuit even though he and his lawyers cannot provide more details about what caused the plaintiff's injury. The Rules thus make a conscious choice to let plaintiffs sue when they cannot be specific about major issues in the case, even though in some cases the defendant will have to hire a lawyer and spend time to rebut a case that turns out to have no basis. To many defendants, that tradeoff seems unfair, but the counterargument is that it is more unfair to permit a defendant to avoid liability because the plaintiff cannot learn enough to file a case if the defendant has the key evidence under its control.

There are a few situations in which there is a different tradeoff. Rule 9(b) requires that fraud or mistake be pled "with particularity," but certain other arguably similar states of mind, such as malice, need not. The rationale behind the fraud exception sometimes is stated to be based on reputational harm, but since a complaint may allege, without particularity, that the defendant is a liar, murderer, or child molester, that supposed rationale for Rule 9(b) is hard to defend. It would also not explain why Rule 9(b) also requires that a complaint alleging even an innocent "mistake" must have additional specificity. The better view is that a claim of fraud is too general and does not

[37] The continued viability of those tradeoffs in light of recent Supreme Court decisions interpreting Rule 8 is discussed *infra*.

give the defendant sufficient information to begin to prepare a defense, which is why particularity is required for claims of fraud and mistake. In those specific kinds of cases, the rationales for allowing notice pleading are overcome by the competing needs of a defendant to be able to start a defense, and hence a different tradeoff was made. But in both the general rule and the exception, one cannot understand what the Rules are doing unless one understands the reasons behind both Rules and the tradeoffs that they embody.

The second tradeoff relates to the first but is slightly different in effect and results. It can most clearly be illustrated by focusing on a case where fraud is alleged, and hence the complaint must be specific as to whether, for example, the defendant's sale of its stock was fraudulent because it failed to include certain expenses in its income statement or it knowingly overvalued accounts receivable on its balance sheet. The question still remains, how much evidence (admissible or otherwise) must the plaintiff have, at the time he files the complaint, in order to support his allegations of fraud? In general, the Rules allow a plaintiff, or more accurately his lawyer, to make such allegations so long as he has a good-faith belief that the facts are true and that proof can be obtained through discovery.[38] In short, not very much is required, which many defendants argue is an ill-advised tradeoff. Whatever one thinks of that tradeoff, there is no doubt that it helps plaintiffs stay in court until they can take discovery, a result that a higher level of required prefiling proof would not allow.

This tradeoff was, for a period in the 1980s and early 1990s, different when Rule 11 was tightened in response to alleged abuses by counsel for plaintiffs in bringing frivolous cases. After a relatively brief experiment, the Rule was changed back to close to its original balance, largely because the alternative tradeoff was thought to unduly discourage meritorious litigation by placing counsel and clients at risk of being forced to pay heavy monetary sanctions. There was also another undesirable side effect of the stricter sanctions Rule: it led to extensive collateral litigation over whether a pleading lacked a reasonable factual or legal basis, which many judges found unpleasant, time-consuming, and often counterproductive. This example also illustrates the proposition that, in designing rules, it is important to consider their administrability and to be wary of theoretically perfect rules that necessitate extensive litigation to carry out, especially when the additional litigation is unrelated to the merits.

---

[38] FED. R. CIV. P. 11(b)(3).

As we will see below, sometimes the Rules prefer fairness (and complexity) over ease of administration, but not always.

Two recent Supreme Court decisions have called into question whether the balance that generally was thought to have been struck in Rule 8, favoring allowing cases to proceed beyond the motion-to-dismiss stage unless the plaintiff's claim was insufficient as a matter of law, continues to exist, although the text of Rule 8 remains unchanged. Without exploring their rationales in detail, or debating whether the Court's interpretations of Rule 8 were proper, there can be little doubt that the decisions in *Bell Atlantic Corp. v. Twombly*[39] and *Ashcroft v. Iqbal*[40] imposed a pleading standard for complaints that gave greater weight to the interests of defendants in avoiding discovery and in having to defend against the plaintiff's claims than had previously been the law. Quite apart from the question of whether the balance struck in those cases is the proper one, there is also a substantial issue as to whether the Court should have, in effect, reset the balance in Rule 8 by reinterpreting the Rule, or do what it has done in the past when similar pleas were made: relegate the issue to the Civil Rules Committee to do the job.

The Rules Committee has begun an examination of Rule 8, and Congress has also considered restoring what some considered to be the status quo ante pre-*Twombly* while the Committee decides what the appropriate tradeoff should be and whether Rule 8 is the place to make any adjustment. There are many possible alternatives to both the prior and the current understandings of Rule 8, and the Rules Committee is best equipped to evaluate where the proper balance lies and how to craft a rule reflecting the desired tradeoff. By way of illustration, I have included as an Appendix a proposal under which a plaintiff would be given the opportunity to present his claim to the putative defendant before filing suit, in effect as a means of testing the waters. If the defendant did not provide a basis for contesting the plaintiff's factual allegations, the plaintiff's burden of pleading in the complaint would be minimal. However, if the defendant chose to disclose its version of the facts (and perhaps the law), the plaintiff would have to include more in the complaint to survive a motion to dismiss. The proposal embodies a series of tradeoffs that are different from those under current law, but it is hopefully a means of increasing

---

[39] Bell Atl. Corp. v. Twombly, 550 U.S. 544 (2007).

[40] Ashcroft v. Iqbal, 129 S. Ct. 1937 (2009).

the fairness of the process to plaintiffs, as they see the law after *Iqbal*, and to defendants, as they saw it before *Twombly*.

B. Service of Pleadings

Filing a complaint is the first step in a federal court case; after that, the defendant must be served. Leaving aside the due process and statutory questions about when a court may exercise personal jurisdiction over a defendant, the rules also must establish the proper means to effect service of the complaint. Again, assuming no constitutional problems with the particular method of providing notice, the choice involves tradeoffs between competing values in at least two respects: who may accept service on behalf of the defendant, and who is eligible to make that service?

If the defendant is an individual, is service on the defendant's spouse at the defendant's residence sufficient? What about on one of the defendant's children, or might the answer depend on the age of the child? What about others residing with the defendant, or his co-workers? Might the nature of the suit matter, or the amount in controversy, or would such differences lead to disputes over peripheral issues, what I call "side shows"? Suppose the defendant is a corporation or an agency of government that can act only through individuals: must the CEO or agency head be the one who receives the complaint, and if not, should the rules take into account the possibility that the complaint will not be delivered to the appropriate person within the entity?

As for the person making service, at one time the Federal Rules required that someone in the office of the U.S. Marshal for the district in which the case was filed conduct all service of process by personally delivering the summons and complaint to the defendant or other person authorized to accept service.[41] That was thought necessary so that there would be no disputes about whether the certificate of service was truthful, but it also proved very costly and became nearly impossible to continue as the federal court civil caseloads increased from 34,734 civil filings in 1940—shortly after the Rules went into effect—to 168,789 in 1980[42] when the predecessor of current Rule 4(c)(2) was changed to allow any person

---

[41] FED. R. CIV. P. 4(c) advisory committee's note (1980) (suggesting that the requirement for service by the Marshal was not absolute and deciding to make clear that it is only one option).

[42] ADMIN. OFFICE OF THE U.S. COURTS, 1980 ANNUAL REPORT OF THE DIRECTOR 3 tbl.3 (1980).

over the age of eighteen, not a party to the case, to effect service. Did that mean that there were no post-1980 disputes over service of process? Of course not. But the drafters were willing to accept some additional litigation as a tradeoff for the reduction in cost of service of process and the delays that ensued when only the Marshal's office was permitted to do the job. In addition, service by mail is acceptable in some circumstances now,[43] not because the mail system is perfect, but because the risk of nondelivery is relatively small, it can be guarded against by other means, and the convenience and lower cost are thought to be worth the tradeoff in certainty. And once the parties have entered their appearances through counsel in the case, service by mail (and now e-mail) is the norm for similar reasons of cost and efficiency.

Similar tradeoffs of efficiency versus certainty also apply when it comes to who is a proper person to accept service for an individual or an entity. Rule 4(e)(2)(B) allows the person making service to leave the complaint with a person of suitable age at the home of the defendant on the theory that such persons will generally see that the actual defendant receives it promptly. The contrary rule would result in added costs of attempting service and in delays in moving the case forward. Those might initially be borne by the plaintiff, but it would also be possible to shift those costs to the defendant if the plaintiff prevails. The tradeoff that allows service on others who live with the defendant has not been extended to allow service when the recipient is not a resident of the home but simply a coworker. However, when the complaint relates to the work of a business or government agency, other practical rules apply, and service can be made in ways that provide reasonable but less-than-total assurance that the persons who need to see the complaint will receive it. These kinds of tradeoffs differ from those made in creating Rule 8, governing the contents of a complaint, because they do not make it easier for one side to remain in court and perhaps even to prevail. Rather, they are part of an overall effort to reduce cost and delay to all parties, notwithstanding a small risk of error that is considered to be an acceptable tradeoff for the offsetting benefits.

---

43 FED. R. CIV. P. 4(d)(1)(G); *id.* 4(e)(1) (allowing service by mail in accordance with state law); *id.* 4(i) (explaining when service on the United States and its officers and agencies may be made by registered or certified mail).

C. Defendant's Response Options

Generally, most defendants have twenty-one days after being served with the summons and complaint in which to answer or move to dismiss the complaint on one of the grounds set forth in Rule 12(b).[44] The Rules could have given defendants a "reasonable time" to reply, which is what happens in the real world becase motions for extension are routinely agreed to by the plaintiff's counsel or granted by the court. But having a fixed time in the Rules eliminates some but not all uncertainty, as evidenced by the complexity of Rule 6 on computing time. Given the ease with which extensions are granted, the actual number of days is not crucial, because the Rules simply set a default time period and give an indication of what might be considered reasonable—in contrast to five days or five months.

As noted above, plaintiffs do not have to provide any "law" in their complaint, let alone cite relevant cases to support their claim. If that seems unfair to defendants, Rule 12(b) gives them a chance to go on the attack and ask the court to dismiss the case on any of several grounds, the most significant one being that the plaintiff fails to state a claim on which relief may be granted. In plain terms, this means that the plaintiff has no legal claim based on the facts set forth in the complaint. If the defendant contends that the plaintiff has no viable legal theory, the defendant must accept as true all of the facts alleged in the complaint. Thus, for example, in an auto accident case, where the claim is that the defendant passenger distracted the driver by talking to him and caused the driver to hit the car in which the plaintiff was riding, the passenger could move to dismiss that claim on the ground that the law does not make a passenger liable to a third party even if the alleged distraction was a factual cause of the accident. Alternatively, the defendant might move to dismiss the complaint on the ground that the accident took place three years ago and all such cases have to be filed within two years from the date of the accident or they are barred. But the defendant could not move to dismiss on the ground that he did nothing to distract the driver because that dispute is factual rather than legal.

In responding to such a motion, the plaintiff cannot simply say that she has not yet had a chance to prove her case or had time to do all the necessary legal research, but must answer the defendant's legal arguments, often with just a modest extension of time beyond the fourteen days allowed in the Rules. In some cases, the basis of the

---

[44] *Id.* 12(a)(1)(A)(i).

motion will be that the plaintiff failed to allege an essential element of the claim—for example, a federal antitrust violation must involve interstate commerce.[45] If the court agrees, it will generally allow the plaintiff an opportunity to remedy defects of that kind by filing an amended complaint, another practice that helps a plaintiff stay in court and have a chance to prove her case. If an amended complaint is filed, the defendant can once again move to dismiss, and if that motion is granted, the plaintiff will generally not be given a further chance to amend.

Many defendants do not consider the opportunity to file a motion to dismiss an adequate tradeoff for having to respond to a complaint that they consider meritless. However, it at least gives defendants some chance of getting rid of a case early in the process, before the expenses of discovery are incurred. And whether it is in fact adequate is of less significance than is the fact that the tradeoff was consciously made based on a desire to see that plaintiffs are given a reasonable opportunity to establish the merits of their claim, including taking discovery if they are shown to have a valid legal claim but need to gather factual support for it.

D. Attorneys' Fees

Another essential element of our civil justice system involves two rules relating to attorneys' fees. Although not part of the *Federal Rules of Civil Procedure*, they are at least as significant as the notice concept for complaints in making it possible for ordinary people to pursue civil claims for money damages. The first is called the American Rule, under which each side bears the costs of paying its own attorney, as well as many but not all costs of litigation, win or lose.[46] The reason that this matters so much is that a plaintiff who has a claim against a well-financed defendant may be willing to risk losing the case, and perhaps having to pay his own lawyer, but the prospect of having to pay for defense counsel as well, which is generally the law in England,[47] could make the risk so great that the case is not worth bringing. This pro-plaintiff rule also means that, if some dubious cases are brought, defendants will have to pay their

---

[45] *See, e.g.*, Sherman Antitrust Act, 15 U.S.C. §§ 1–7 (1890).

[46] BLACK'S LAW DICTIONARY 98 (9th ed. 2009).

[47] *Id.* at 609 (defining the English rule, under which "a losing litigant must pay the winner's costs and attorney's fees").

own lawyers to get them dismissed. The tradeoff is nonetheless considered, on balance, better than having valid claims not brought because of a fear of having to pay counsel fees for the defendant. Indeed, both Congress and state legislatures have passed laws that require losing defendants, but generally not losing plaintiffs, to pay the fees of their opponents where, for reasons of public policy, the legislature wished to make bringing certain kinds of lawsuits—the most significant being claims for civil rights violations—more attractive, and hence changed the tradeoff to create a one-way shifting of attorneys' fees for those cases.[48]

The second rule allows a client to agree to pay a lawyer only if the client wins or obtains a favorable settlement, and to use the resulting award for that payment. Known as "contingent fees," they are, in conjunction with the American Rule, a major assist to plaintiffs who wish to bring a lawsuit, often for personal injury claims, but cannot afford to pay their lawyer if they lose. There is at least one exception: contingent fees are generally not permitted in criminal cases and in matrimonial matters, on the theory that the potential for misconduct in those cases creates improper incentives, thus making the tradeoff undesirable from a public policy perspective.[49] In the real world, those restrictions are evaded by entering fixed-fee agreements, which are realistically capable of being fulfilled only if the defendant is acquitted (or receives a very light sentence) or the client achieves a favorable resolution in the divorce. There have been various proposals to modify these rules, generally to make them more favorable to defendants, but there is little likelihood that the current tradeoffs will be changed in any significant respect in the United States in the near future.

E. Rule 11

The previous Subparts discussed the tradeoffs that generally favor the plaintiffs both in allowing complaints to go forward without actual proof of the factual allegations and in not requiring losing plaintiffs to pay the fees of a prevailing defendant. Rule 11, which allows courts to order payment of the other side's attorneys' fees in some situations, is a modest counterweight to that balancing of interests. Of equal significance is that its three-part history shows how the tradeoffs in this area have been significantly recalibrated from time to time, in

---

[48] *See, e.g.*, 42 U.S.C. § 1988 (2006).

[49] MODEL RULES OF PROF'L CONDUCT R. 1.5(d) (2011).

part based on less-than-satisfactory experience with the Rule as then in effect.[50]

As issued in 1938, Rule 11 was largely hortatory: attorneys should stop, look, and listen before filing pleadings, but nothing was done if they failed in this duty—except in what everyone would agree are the most egregious situations. Objections were raised, almost entirely by defense counsel and their clients, that the Rule should be amended to add teeth in order to discourage frivolous litigation, and in 1983 the Supreme Court approved amendments that radically altered the role of Rule 11. Without going into great detail, the major changes were: (a) sanctions became mandatory if a violation was found, with attorneys' fees to the other side as the preferred remedy; (b) clients, as well as lawyers, could be held responsible for unwarranted assertions of law; and (c) lawyers were required to more closely question their clients on assertions of fact,[51] such that they almost became required to take an adversary position to them. Sanction motions routinely accompanied motions to dismiss, and the threat of seeking sanctions was often employed by defendants. Courts were often required to hold sanction hearings and to determine appropriate fees to be paid by both clients and lawyers. In theory, defendants and their lawyers could have been held liable under the Rule, but almost all of the activity was aimed at plaintiffs, especially in civil rights cases. A veritable cottage industry developed on Rule 11, but it also stimulated very substantial backlash, including from judges who found the Rule to be burdensome and distracting from their main duties.

Ten years later, the Court largely reversed itself and brought Rule 11 back much closer to the 1938 version than to the 1983 edition. Sanctions were no longer mandatory even if a violation was found; a twenty-one-day safe harbor was created to allow a party to amend or withdraw a pleading in that time, without incurring the risk of sanctions;[52] attorneys' fees were no longer the remedy of choice;[53] and clients were no longer held accountable for the bad legal advice that their lawyers gave them.[54] In addition, factual allegations can be

---

[50] Sanctions, including attorneys' fees, are also available under Rule 37(b) in discovery disputes; 28 U.S.C. § 1927 is also a source of fee shifting in some circumstances. For purposes of this Article, Rule 11 illustrates the point in sufficient detail without discussing other sources of fee shifting and the tradeoffs they entail.

[51] FED. R. CIV. P. 11 (1983).

[52] FED. R. CIV. P. 11(c)(2) (1993).

[53] *Id.* 11(c)(4)–(5).

[54] *Id.* 11(b)(2) (only counsel and unrepresented parties subject to this requirement).

safely made "on information and belief" if the lawyer identified them as such, provided that the lawyer reasonably believed that those allegations "will likely have evidentiary support after a reasonable opportunity for further investigation or discovery."[55]

This three-part history of Rule 11 illustrates both the differing tradeoffs that are possible in dealing with the problem of what level of pre-suit investigation a plaintiff and her counsel must make, and how judgments as to the appropriate balance changed over time and with experience under other options. There are still some, again mainly on the defense side, who are unhappy with the balance, but there is no significant movement to use Rule 11 as an additional counterweight to the notice-pleading concept or the American Rule on attorneys' fees. The tradeoffs in Rule 11 seem in equipoise at least for the time being.

F. Discovery

It may be difficult for civil litigators today to realize that, until the Federal Rules became the law in 1938, there was very little discovery in most civil cases in the United States. Although perhaps never spelled out fully, the rationales for only a limited discovery regime include the reality that discovery costs money and causes delays. To some, the concept of discovery runs counter to the notion that each side in the adversary system should prepare its own case and not be required to help the opponent. The Federal Rules make a different set of tradeoffs: justice is served by having the truth come out at trial, and that is best accomplished by full pretrial discovery of what each side knows, which may also lead to earlier, or at least better-informed, settlements. That choice is not without downsides, which include significant increases in costs and new (and sometimes excessive) burdens on the party who is on the receiving end of some discovery requests. The *Federal Rules of Criminal Procedure*, also adopted by the Supreme Court, make a quite different judgment about the desirability of pretrial discovery.[56] Under them, discovery is quite limited, subject to the constitutional requirement—itself a form of tradeoff—that the prosecutor furnish the defendant with exculpatory evidence in his possession.[57] Furthermore, civil justice systems in other countries (and in most states for many years after the *Federal*

---

[55] *Id.* 11(b)(3).

[56] FED. R. CRIM. P. 16.

[57] Brady v. Maryland, 373 U.S. 83 (1963).

Rules of Civil Procedure were adopted, but not today) have made different tradeoffs and do not have anywhere near the kind of extensive discovery allowed under our Federal Rules.

Within this broad pro-discovery regime, there are various provisions that alter the balance, at least to some degree. Until 1970, Rule 34 required parties to seek court approval for requests for the production of documents from another party, although parties often agreed to produce at least some documents without a court order. The apparent rationale under the original Rule was that examining the opponent's records was a highly invasive process, unlike answering interrogatories or having one of your witnesses examined orally at a deposition, and hence required court supervision. But as document production became more common, and judges were burdened and costs were increased because of the necessity of filing routine motions to produce documents, Rule 34 was changed to require the person on whom the request was served to answer it or to object, with the requesting party then being forced to file a motion if it wished to challenge the objection. Similarly, under Rule 45, until 1991, a party wishing to obtain documents from a third party could do so only by serving a subpoena on that third party with a notice to take an oral deposition and a request that the deponent bring certain documents.[58] The Rule has been changed so that a third-party subpoena can now seek only documents without the necessity of having the custodian show up and present them—a recognition that burdens and costs needed to be adjusted in light of current litigation practices.

Several other aspects of discovery involve somewhat different tradeoffs. The principle of full discovery was tested early on in *Hickman v. Taylor*.[59] Plaintiffs requested a wide range of documents from the defendant, including copies of witness statements obtained by counsel (or investigators hired by counsel), counsel's notes of meetings with witnesses, and summaries of oral meetings with witnesses for which there were no statements or notes.[60] The Supreme Court ruled that the request was for a category of documents that the circuit court labeled the "work product of the lawyer" and held that, because the plaintiffs could interview or depose those witnesses, they should not be allowed, in effect, to freeload on the work of the defendant's counsel and, perhaps more importantly, gain

---

58 FED. R. CIV. P. 45 (1990) (amended 1991).

59 329 U.S. 495 (1947).

60 *Id.* at 498–99.

access to their thought processes by the use of this form of discovery, absent some special need or justification, such as the death of a witness.[61] To reach that result, the Court construed the Rules, which were completely silent on this issue, to create an implied privilege, with a different set of tradeoffs than the general rule of broad discovery—a result that has since been codified in Rule 26(b)(3). Like the original Court opinion, that Rule is not absolute but allows discovery of some work product if there is a showing of substantial need and undue hardship in obtaining the information elsewhere. Even then the court is required to protect against revealing the lawyer's mental processes regarding the case.

A third example of a different implicit tradeoff relates to discovery involving a party's expert witnesses. Without some special rule, the plaintiff could ask the defendant the name of any expert that the defendant had consulted regarding the case, and then the plaintiff could take that person's deposition, paying only the normal minimal witness fee owed to ordinary fact witnesses, and do the same for every expert who was consulted. The Rules, including recent amendments that make somewhat different tradeoffs, now limit discovery to experts a party has identified as potential trial witnesses, and even they can be deposed only after they have submitted a report concerning their proposed testimony.[62] The expert witness is entitled to be paid at his or her commercial rates, for preparation and for attending the deposition, by the party taking the deposition.[63]

Finally, there is the tradeoff on the scope of discovery. It is clear that discovery is not objectionable because it seeks evidence that may be inadmissible, so long as it "appears reasonably calculated to lead to the discovery of admissible evidence."[64] That limitation is quite modest, and parties, especially in the most significant commercial cases, began to request extremely broad and burdensome discovery. Because the requesting party did not know what would be produced,

---

[61] *Id.* at 509–11.

[62] FED. R. CIV. P. 26(b)(4)(A) (2010).

[63] *Id.* 26(b)(4)(E)(i)–(ii). Another recent amendment added Rules 26(b)(4)(B)–(C), under which communications between counsel and their testifying experts are generally not discoverable (except for information about fee arrangements and a few other limited subjects). The reason for the limitation is that the present system was found to be inefficient and, for lawyers who knew the system, easy to evade by not putting discussions in writing. This and another change in Rule 26(a)(2)(C) regarding discovery of witnesses who provided both factual and expert testimony are further indications of the ever-shifting nature of the various tradeoffs in the rules on discovery from expert witnesses.

[64] *Id.* 26(b)(1).

it was impossible to know in advance whether it would produce relevant information. In addition, not all information that is technically relevant is of equal significance in a case, and the burden of producing some documents may be much greater than producing others.

One answer that the Rules provide is to require that information sought be relevant to the "party's claim or defense" in the case, while allowing discovery as to the "subject matter" only with court approval and for "good cause."[65] More significantly, what is now Rule 26(b)(2)(C) made explicit the power of the court to limit discovery that was unduly burdensome or cumulative by adding an explicit cost-benefit analysis that allows the judge to decide when enough is enough.[66] This rebalancing reflects a different calibration in the tradeoffs between full discovery and other values, informed by the way that current litigation is being conducted. And most recently, the sea change that electronic record keeping has made caused the drafters to add provisions calling special attention to discovery of electronic records and the need to balance the benefits and burdens of that subset of records with its special characteristics regarding search and retrieval capabilities in mind.[67]

III

JOINDER OF CLAIMS AND PARTIES

A. Traditional Lawsuits

One innovation of the Federal Rules is the recognition that there should be procedures available to enable the parties to add related claims and to enable and, in some cases, require that third parties be brought into what is a traditional case between a single plaintiff and a single defendant. That innovation mainly removed existing barriers to that type of joinder and, in most of its manifestations, involved few if any significant tradeoffs. There are, however, three Rules that do require tradeoffs and that are worthy of note because the beneficiaries of those Rules are often not the existing parties.

Rule 13 deals with counterclaims, which are pleadings in which a defendant asserts a claim against the plaintiff that seeks affirmative relief for the defendant. Rule 13(a) provides that, if a defendant has a

---

65 *Id.*

66 *Id.* 26(b)(2)(C)(i)–(iii).

67 *Id.* 16(b)(3)(B)(iii); *id.* 26(f)(3)(C); *id.* 34(a)(1)(A), (b)(2)(D)–(E); *id.* 37(e).

counterclaim that arises out of the same transaction or occurrence as the main claim (for example, a single automobile accident), the counterclaim is "compulsory," meaning that, with limited exceptions, the defendant must assert it in this case or be barred from bringing it in the future. In most situations, the defendant will want to add the counterclaim, but in some cases the defendant might prefer to wait, or to bring it in another forum, but Rule 13(a) forbids that. Beyond whatever interest the parties have in seeing that the present case includes all related claims, there is an independent interest of a variety of third parties—including witnesses, judges, jurors, and other litigants who are waiting to have their cases heard—in seeing that the events giving rise to this lawsuit are the subject of only one judicial proceeding. Through the compulsory counterclaim rule, the interest of the defendant in choosing the time and place of filing its claim is subordinated to the overall efficiency interest of the judicial system in having only one lawsuit for the controversy.

Similar tradeoffs are made under Rule 19 (compulsory joinder of parties) and Rule 24 (intervention). Under Rule 19, either an existing party or the court can point to the absence of a third party who is connected to the main transaction and who ought to be joined, if possible, in the interest of fairness or efficiency (although those are not the terms actually used). If the requirements of the Rule are met, the party will be added as a plaintiff or defendant, as appropriate, even if the existing parties oppose the joinder. However, if joinder is not possible because it will destroy subject matter jurisdiction or because the court lacks personal jurisdiction over the party to be joined, the court must choose between allowing the action to proceed without joinder and dismissing the action, which will generally occur only if there is another forum where all parties can be joined. The subject matter jurisdiction limit is significant because § 1367(b) precludes the use of supplementary jurisdiction to bring in additional nondiverse parties under Rule 19,[68] in effect reaffirming the terms of the tradeoff made by Rule 19.

Rule 24, which is often seen as a companion to Rule 19, allows third parties to seek to intervene in cases in which neither the plaintiff nor the defendant wishes to have them involved, but in which they have a substantial interest. The conditions under which intervention, either general or limited, will be granted are not significant for this Article. What is important is that the Rule allows the interests of third

[68] 28 U.S.C. § 1367(b) (2006).

parties and the systemic interest in overall efficiency to trump the interest of the existing parties in confining their lawsuit to the parties and issues they have chosen.

B. Class Actions

The 1938 version of Rule 23 allowed some class actions, but it was not until 1966 that the Rule had real teeth. The big change was the creation of the damages class action in Rule 23(b)(3), and in particular the decision to allow classes to go forward without class members having to take an affirmative step to join the class, but having the right to exclude themselves from the class—to "opt out." Given the inertia among class members, especially in cases with small claims, insisting on "opt-in" classes would mean that most class actions would not go forward, which is plainly what defendants prefer. The advantage given the plaintiffs by permitting opt-out classes can be seen as part of the overall effort to provide meaningful enforcement for federal and state law in the federal courts, and not just have laws on the books that do little to help their intended beneficiaries. But at the same time, it substantially increased the stakes for defendants sued under Rule 23.

The tradeoff was tempered by certain burdens that were placed on the named plaintiffs or, more realistically, their lawyers in damages class actions. The first and most important was the requirement that, upon class certification, personal notice (which generally means notice by mail) had to be provided to each class member who could reasonably be identified and that the class (counsel) had to pay for it.[69] Due process surely requires some kind of notice in order to bind the class, but where the cost of notice exceeds the value of the claim, or even amounts to a significant fraction of it, the requirement of individual notice can only be seen as a tradeoff that benefits defendants. Indeed, it was defendants who insisted that the Rule be strictly followed, and the Supreme Court agreed.[70]

The second major tradeoff in class actions, again motivated in part by due process considerations, relates to settlement. Except for cases involving minors or persons who are adjudicated incompetents, parties are free to settle non-class cases without court approval. But Rule 23(e) changes that, by requiring notice to the class, a hearing,

---

[69] Eisen v. Carlisle & Jacquelin, 417 U.S. 156, 176 (1974).

[70] *Id.* at 178–79.

and a specific finding by the court that any settlement of a certified class and any fees paid to class counsel be reasonable, regardless of what the named plaintiffs have agreed to accept and defendants have agreed to pay.[71] The basic theory behind this requirement is that, because the entire class will be bound (unless a class member opts out), the possibility of selling out the class and overpaying class counsel necessitates some outside supervision because class members cannot be expected to monitor the case and they lack the power to stop improper settlements without assistance of the court.

IV

ERIE AND STATE LAW

The statutes granting jurisdiction over diversity cases do not establish what law should apply in those cases. Because there is no federal statute, treaty, or constitutional provision on which the claim is based, the federal courts could, at least in theory, apply federal common law, much as states apply their own common law for cases in which there is no applicable substantive statute. That was the practice for nearly one hundred years, until the Supreme Court ended it in *Erie Railroad Co. v. Tompkins*,[72] in which it held that state law, both common and statutory, governed the substantive aspects of diversity cases in federal court. In part, *Erie* is a judicial recognition that policy choices and tradeoffs are made through state common-law adjudications as well as by the legislatures, and that those choices should be recognized by the federal courts absent specific federal law to the contrary. In the same year that *Erie* was decided, the *Federal Rules of Civil Procedure* became effective. Most of those Rules were clearly procedural, not substantive, but there was still a question in some cases as to which law would govern if there were an otherwise-applicable state law.

The issue was starkly presented in the statute-of-limitations case of *Hanna v. Plumer*.[73] A Massachusetts law required that a complaint be served personally on the executor of an estate, whereas the federal rule allowed the summons and complaint to be left with a person of suitable age at the home of the defendant. The Court had previously held that state statutes of limitation were substantive for *Erie* purposes, and Massachusetts treated its service-of-process rules as

---

71 FED. R. CIV. P. 23(e)(1)–(2), (h).

72 304 U.S. 64 (1938).

73 380 U.S. 460 (1965).

part of its statute of limitations. There was no dispute that the executor actually received the complaint after it was left at his home, but because the claim arose under state law, the executor argued that the service rules were also substantive and hence state law, not the Federal Rule, controlled.

The Court rejected the executor's claim and applied the Federal Rule, which meant that the statute of limitations had been satisfied. In subsequent *Erie* cases, a number of which involved statutes of limitations, the Court has generally held that the Rule controlled, but not always. This is not the place to assess the merits of the Court's approach generally or as applied to particular cases, except as it relates to whether those cases illustrate the tradeoff principle. At least in theory, the Court could have avoided the case-by-case resolution of tradeoffs it appears to have made under *Erie*. One argument often made in these cases is that uniformity is a vital concern, and so the Federal Rules should always trump contrary state law if any Federal Rule is even arguably applicable. The Court has recognized the importance of uniformity—*Hanna* is an example of where uniformity was a significant reason to support the result—but in some cases it has allowed local law that had a strong procedural element to it, such as *Gasperini v. Center for Humanities, Inc.*,[74] to govern at the price of loss of certainty and uniformity. It has done so when important state policies were at issue, which justified the necessary tradeoff of rendering a Federal Rule at least partially inapplicable. To be sure, in all *Erie* cases there are federal statutes that have considerable impact on the Court's decisions, but none of them is so clear that they deny the Court any room for interpretation. Thus, in construing these statutes and its own prior decisions in cases arising under state law, the Court has made, sometimes only implicitly, tradeoffs between uniform federal procedural law and upholding state policy choices in areas that are close to the procedural line.

V

APPEALS

There are three examples applicable to appeals in the federal courts that involve clear tradeoffs. The first is the final judgment rule,[75] under which an appeal may only be taken from a final judgment, a

[74] 518 U.S. 415 (1996).

[75] 28 U.S.C. § 1291 (2006).

term that is fairly strictly construed to mean only when the case is finally concluded. That approach involves very significant tradeoffs (for example, denying an immediate appeal when a defendant's motion to dismiss a case for failure to state a valid legal claim is rejected), thereby triggering extensive discovery and perhaps even a trial. In such cases, defendants believe that the final judgment rule imposes great burdens on them and the trial court because an immediate appeal might end the case. On the other side, plaintiffs will object to immediate appeals because they will delay the outcome or may require the plaintiff to litigate in both the trial and appeals courts at the same time. Because most cases settle (although with a different balance when the plaintiff has successfully resisted an unappealable motion to dismiss), the refusal to allow an immediate appeal is not just a postponement of the appeal, but may result in no appeal at all. The final judgment rule is clearly a tradeoff, rejecting the interest of the would-be appellant in an immediate appeal in favor of an appeal only at the conclusion of the case, at which time all issues remaining in the case can be taken up.

The tradeoff embodied in the final judgment rule is not universally accepted. New York, for example, allows a wide range of interlocutory appeals,[76] and other jurisdictions fall somewhere in between. Indeed, even in the federal system, the final judgment rule is not an absolute, but has exceptions whose underlying theme is that some decisions are so important that failure to allow interlocutory review will place unreasonable burdens on the parties or the court.

One set of these exceptions is embodied in the collateral order rule, under which a narrow set of orders that plainly do not resolve the entire case are reviewable because of their importance and because, as a practical matter, if review is postponed until the end of the case, a reversal of the decision will not vindicate many of the policies behind the rule that the appellant urges.[77] The most significant category of those cases involves rejections of claims by government officials for various kinds of immunity, often from liability for claims for money damages for alleged violations of constitutional rights. The Court has held that such immunities do not merely remove the official from any liability, but they relieve that person from having to defend the case at

---

[76] N.Y. C.P.L.R. § 5701 (McKinney 2011).

[77] Cohen v. Beneficial Indus. Loan Corp., 337 U.S. 541, 546 (1949). Recently, the Court unanimously rejected a claim that the denial of a claim of attorney–client privilege fell within the collateral order rule and denied the right to an interlocutory appeal. Mohawk Indus., Inc. v. Carpenter, 130 S. Ct. 599, 608 (2009).

all.[78] The extent of this exception to, or perhaps more properly, an interpretation of, the final judgment rule is not significant for these purposes. Rather, what is important is that the exception represents a different balancing of the relevant interests in this category of cases, resulting in the official obtaining a right to an interlocutory appeal not enjoyed by most defendants.

There are other examples of a right to an interlocutory appeal, such as from an order granting or denying a preliminary injunction, where an appeal only from a final order may be moot as a practical matter. But most of the exceptions make the appeal discretionary, unlike an appeal from a final judgment, which is a right. For example, most denials of motions to dismiss or motions for summary judgment are not appealable, but 28 U.S.C. § 1292(b) allows a district court to certify such an order for immediate appeal under certain circumstances and allows the court of appeals to agree to hear that appeal if it chooses to do so. Similarly, a motion granting or denying class certification has, since 1998**,** been considered so crucial to the outcome of the case that the losing party may seek immediate review of the class certification ruling under Rule 23(f), although the court of appeals has discretion as to whether to hear it. In short, although the tradeoff between an immediate appeal and awaiting a final judgment is generally resolved in favor of the latter, that is not true in some categories of orders, where the desirability of an early appeal is seen to outweigh the normally overriding considerations to the contrary.

A second example of tradeoffs in the appeal area involves the time for taking an appeal, or, more precisely, the fact that, unlike most litigation deadlines, that time cannot be extended except in narrowly defined circumstances. The basic time is thirty days from entry of the final judgment,[79] but that period is automatically extended if a motion is made under either Rule 52 or 59 to set aside the judgment. Such a motion must be made within twenty-eight days from entry of judgment[80]—and, unlike almost every other civil motion—that period, which was ten days until December 1, 2009, cannot be extended.[81] Once the judgment becomes final, the thirty days for filing an appeal can be extended only for very limited and quite specific reasons, generally those for which the appellant is not

---

78 Mitchell v. Forsyth, 472 U.S. 511, 530 (1985).

79 FED. R. APP. P. 4(a)(1).

80 FED. R. CIV. P. 52(b), 59(b).

81 *Id.* 6(b)(2).

responsible, such as the clerk not sending out the final order.[82] But even then, the forgiveness is limited in duration, and the terms of the exception are narrowly confined.

Most recently, the U.S. Supreme Court found an appeal untimely, even though the appellant had acted in reliance on an error made by a federal judge and that error was not objected to, or even noted by, the opposing party, causing his appeal to be filed two days late.[83] In the Court's view, the times for taking an appeal are "jurisdictional," meaning that neither the parties nor the courts can consent to or even order their enlargement, presumably because the virtues of certainty and finality outweigh the interest in fairness that a more flexible approach would countenance. There is, of course, nothing in the nature of a time limit for an appeal that is any more fixed than any other deadline, and Congress could, if it wished, change the statutes, or the Court could change some of the Rules to produce a different tradeoff. Those who disagree with the rigidity of the Supreme Court's reading of the governing authorities do not suggest that rules about timing do not inevitably involve tradeoffs of various kinds; it is that they disagree with the choices the Court ascribes to those who wrote those statutes and rules.

Third, the degree of deference given by appeals courts to decisions of trial courts and juries is another example of a set of tradeoffs. One of those is contained in the Seventh Amendment, which generally forbids the reexamination of facts found by a jury.[84] As a result, there is a very high—some would say nearly conclusive—willingness to tolerate jury error as a lesser evil than having appellate judges substituting their views on the facts for those of the jury. But when the fact finder is the trial judge, Rule 52(a)(6) allows the court of appeals to set aside a factual finding if it is "clearly erroneous," a more rigorous standard of review, under which a single judge's view of the facts is given less deference than is that of a jury of six or more. There are those who would argue that, despite the advantage that a trial judge has in seeing live witnesses and being better able to judge their credibility, three judges with an opportunity to read the full trial record and discuss it with each other are at least as likely to make the correct factual findings, if not more so. Even if, however, an appellate panel were better able to reach the correct result, the same

---

82 FED. R. APP. P. 4(a)(6).

83 Bowles v. Russell, 551 U.S. 205, 207, 214 (2007).

84 U.S. CONST. amend. VII.

standard might continue to be used because changing to a de novo review of factual findings would greatly increase the burdens on appellate courts and would encourage appeals, an alternative tradeoff that is generally considered to be less desirable. And when the issue is whether, for example, the district court should have denied certain discovery or found that evidence that one party sought to have admitted at trial was cumulative, many of the same reasons for limiting review of claimed factual errors, as well as the generally case-specific nature of those questions, support the use of the current "abuse of discretion" standard and the tradeoffs that its use entails.

The added-burden and increase-in-appeals arguments can also be made for issues of law, but there the standard is de novo review. Part of the rationale for that standard is that the trial court has no comparable advantage on legal questions over that gained from seeing witnesses who testified on factual issues. Indeed, the trial judge often has less time to consider legal issues than do appellate court judges, and the briefing at the appellate level is likely to be more complete and focused. Finally, if the trial court makes a mistake on a factual matter, generally only the existing parties will suffer the consequences, whereas if there is an error of law that is not corrected on appeal, that may, as a practical matter, bind many others in similar situations. For each of these standard-of-review issues, different people might make different tradeoffs that would produce different standards, but wherever the balance is struck, there will inevitably be tradeoffs of one kind or another.

CONCLUSION

The idea that procedural rules contain explicit as well as implicit tradeoffs is hardly a novel concept, but it is often one that law students do not appreciate. The primary goal of this Article is to illustrate some of the many ways in which the rules and statutes governing civil procedure inevitably make tradeoffs between competing legitimate objectives. It does not seek to present a comprehensive review of all such tradeoffs, or to evaluate whether the balances struck are correct, or even whether other factors might be at work in reaching them.

This Article's secondary goal is to urge those who write the rules, and the courts that interpret them, to be more explicit in acknowledging the tradeoffs that inevitably must be made. That kind of openness would make it easier to evaluate the balance struck and to

apply the rules to specific cases. Courts in particular become mechanical in some of their interpretations of procedural rules, which is especially unfortunate when there are policy reasons supporting the result that are part of the tradeoff that should be frankly acknowledged. Rarely will a rule be so clear that it admits of only one reading, especially when a case is in a court of appeals, let alone the U.S. Supreme Court. By pointing to the tradeoff supporting the outcome, courts will enlighten the parties and those who have the power to change the statute or rule and will also help lawyers and law students appreciate the inevitable tradeoffs necessary to a well-developed system of civil procedure.

APPENDIX

THE *IQBAL* DILEMMA: A POSSIBLE RESPONSE

Those who represent plaintiffs fear that the legacy of *Iqbal* will be to enable defendants who have exclusive possession of key evidence to obtain dismissal of many complaints because even plaintiffs with a good-faith belief that they can prove their case will not be able to supply the missing allegations that *Iqbal* arguably requires. On the other side, defenders of increased pleading requirements assert that, without some controls, plaintiffs will make unsubstantiated factual allegations and be permitted to embark on extensive—and what defendants consider to be unwarranted and costly—discovery.

One way out of this dilemma is to create an optional pre-suit exhaustion process that will give defendants the opportunity, but not the obligation, to show a would-be plaintiff that there is no factual support for the claim, as a way to dissuade its filing. But if the defendant does not choose to provide that information to the plaintiff, the court, in ruling on a motion to dismiss, could not grant the motion when the missing factual information was under the control of the defendant. I have not attempted to draft a rule embodying this optional exhaustion opportunity, but I have set forth below some examples of how this might have worked in two recent cases, as well as in *Iqbal* and *Twombly*, and then I discuss other aspects of the proposal. Before turning to those cases, I briefly examine a very common type of case in federal court, where the pleadings are always conclusory and sometimes implausible, yet are never dismissed under Rule 12(b)(6): actions under the Freedom of Information Act (FOIA).[85]

A typical and procedurally proper FOIA complaint alleges that the defendant agency has certain records that the plaintiff requested, that the defendant denied the request for the records (and in some cases did not give any factual or legal basis for its denial), and that the denial was without basis in law—that is, none of the enumerated exemptions properly applies. Based on only those bare-bones allegations, the defendant must provide the factual proof for its defense. Even where the agency asserts that the records are properly classified, the courts insist that the agency come forward with some factual basis for its legal claims, no matter how conclusory the

[85] 5 U.S.C. § 552 (2006).

complaint is ("unlawful" or "without basis in law" is pretty conclusory), or how implausible it is that an agency such as the CIA has improperly classified the records, let alone that a court would so find. The reason why those cases are allowed to go forward is that the defendants are in full control of all the key evidence as to the applicability of the exemptions, and it would be unfair to require plaintiffs to show more at this stage of the case, notwithstanding *Iqbal* and *Twombly*. But if the plaintiffs were given access to relevant factual information before filing suit, a court might be justified in insisting that the complaint take that information into account in determining whether Rule 12(b)(6) or perhaps Rule 56 entitles the defendant to dismissal.

One of the areas where *Iqbal* and *Twombly* are expected to have a significant impact is in employment discrimination cases. In a disparate impact case recently decided favorably for the plaintiffs by the U.S. Supreme Court, the plaintiffs proved at trial that a cutoff score used by the City of Chicago to narrow the pool for firefighter applicants had a disproportionate adverse impact on African-American applicants and lacked a business justification, thereby violating Title VII.[86] There was no *Iqbal* problem there, because when Chicago announced the test results, it disclosed the adverse impacts on minorities. But it is quite unlikely after *Iqbal*, especially for private employers, that they would make similar disclosures. The plaintiffs would still file charges with the Equal Employment Opportunity Commission, the defendant would not respond, and a right-to-sue letter would be issued. After *Iqbal*, a plaintiff who alleged only that the cutoff score had a disparate impact on African-Americans and that the cutoff lacked a business justification might have the complaint dismissed for including merely "conclusory" allegations, even though all the detailed information was in the control of the defendant. Under this proposal, if the same sequence were followed, the court would be forbidden from dismissing the complaint on the ground that the allegations were conclusory because the defendant failed to provide the statistics showing the actual impact on the different races and did not offer any evidence, when it had the opportunity to do so before suit was filed.

Or assume that a Toyota suddenly accelerated to ninety miles per hour and crashed into a tree, seriously injuring the driver. Assume the complaint alleged that the car was negligently designed and

[86] Lewis v. City of Chicago, 130 S. Ct. 2191, 2196 (2010).

manufactured and that there were breaches of various warranties, but no additional specifics as to the actual cause of the acceleration. Under at least some readings of *Iqbal*, the defendant might have the case dismissed for failure to include more information as to the cause of the accident. However, under this proposal, if the plaintiff gave Toyota a copy of the proposed complaint containing those allegations, and defendant did not respond, the plaintiff would be able to defeat a motion to dismiss and commence discovery (assuming that there was no other legal basis on which the complaint might be dismissed).

In *Iqbal* itself the Court held that the allegations that the U.S. Attorney General and the Director of the Federal Bureau of Investigation (FBI) approved allegedly discriminatory policies were too indefinite to be allowed to go forward.[87] The plaintiffs had made every reasonable effort to ascertain the facts as to what involvement, if any, those two officials had, including in discovery from other government defendants in that case, but they were rebuffed. Of course, both defendants knew (or at least their official files would show) whether they had approved any policies regarding the detention of aliens after September 11, 2001, of the kind set forth in the complaint; they simply chose not to provide that information. After the case was dismissed against the Attorney General and FBI Director, but allowed to continue against the remaining defendants, the plaintiff settled for the not insignificant sum of $265,000,[88] which suggests that at least some of Iqbal's allegations had considerable merit.

If this proposal had been in effect, the plaintiff would have the option to present the claim to the proposed defendants, either in a letter or a draft complaint, which is essentially what claimants must do when suing the United States under the Federal Tort Claims Act, although there are no similar consequences under that law if the Government remains silent at the administrative level.[89] The defendants would then have a choice: they could ignore the claim, or they could provide sworn statements denying any connection to any policy allegedly covered by the claim, and where appropriate, supporting documentary evidence—for example, copies of orders establishing that the policy was approved by others. If they ignored the claim, they would be precluded from arguing as to information in

---

87 Ashcroft v. Iqbal, 129 S. Ct. 1937, 1954 (2009).

88 Telephone Interview with Alexander Reinert, Plaintiff's Counsel.

89 28 U.S.C. § 2675(a) (2006).

their possession, but not available to the plaintiff, that the complaint failed to set forth the claim with sufficient particularity. They could move to dismiss on the ground that there was no legal basis for the claim, the kind of Rule 12(b)(6) motion that the dissenters and Iqbal's counsel agreed was available. But if the defendant provided relevant pre-suit evidence that countered allegations necessary to establish a claim, the plaintiff would have to present some basis (other than that they disbelieved the defendants) to avoid dismissal—rather like a mini summary judgment. Thus, in the *Iqbal* situation, if the defendants provided only blanket denials of having issued the orders authorizing the challenged detention policies, and did not provide copies of the actual orders that bore on the detention of the class that plaintiff alleged included him, the case could not be dismissed under Rule 12(b)(6). Similarly, if the defendants had not submitted any other evidence that someone else had approved the policies beyond a simple denial that they had not done so, the plaintiff would not be stuck with such unsupported general denials by persons not under oath or subject to cross-examination. In other words, where there is likely to be a paper or e-mail trail, the defendant would generally have to provide the essential parts of it to take advantage of this option.

In *Twombly* the Court found the allegations that the defendants had entered into an anticompetitive agreement to be implausible, especially in light of other allegations that pointed toward conscious, lawful parallel conduct.[90] Under this proposal, if the plaintiff presented the draft complaint to the defendants and they did not respond, the issue of plausibility could no longer be the basis for a motion to dismiss. On the other hand, the defendants could respond with affidavits from senior corporate officials, based on personal knowledge and an investigation described in the affidavits, stating that no meetings on this subject ever took place and that there were no records (paper or electronic) that supported a conclusion that an agreement among the defendants existed. In that case, unless the plaintiff was able to be more specific than the plaintiff was in *Twombly*, the complaint would be dismissed under Rule 12(b)(6).

The advantage of this proposal to plaintiffs, at least compared with the possible negative outcomes under many if not all readings of *Iqbal* and *Twombly*, is that they would have a much better chance of obtaining discovery needed to prove their case. In addition, if the

---

[90] Bell Atl. Corp. v. Twombly, 550 U.S. 544, 566 (2007).

defendants had a valid explanation or defense, the plaintiff (or more precisely the plaintiff's lawyer) would not file the case, for fear of both wasting time and money and possibly suffering Rule 11 sanctions. In addition, in at least some cases, if there were pre-suit exchanges of information in which the defendants would have a stake in conducting them in a meaningful way, settlement discussions based on facts and not just suspicions might occur.

Some of these benefits will also accrue to defendants, but they would have to make a choice when a proposed complaint arrives. In a case with significant financial or other risk, they will surely want to consult with counsel in deciding whether to respond by engaging with the plaintiff. Engaging means some exchange of information, although much less than in formal discovery because the process is voluntary and can be stopped at any point. But there is no reason why this limited discovery cannot run in both directions, so that defendants will have a better idea what to expect and may decide that early settlement is in their best interests as well—something that rarely happens under the current system. And if defendants truly believe that suits are frivolous, telling plaintiffs that early, with evidence to support their position, will help with motions seeking sanctions under Rule 11 or 28 U.S.C. § 1927.

Resorting to such a process might produce statute-of-limitations problems in some cases. Ideally, there would be an automatic tolling of the statute while this option was pending. However, given the limits of the Rules Enacting Act and the Rules of Decision Act, it is doubtful that the rules themselves could accomplish that.[91] Nothing would prevent the parties from entering a tolling agreement, but even if they did not, an exhaustion period lasting only sixty days or so should not cause many statute-of-limitations problems, especially if the process were used promptly after the injury was discovered, before the statute was a real concern.

Plaintiffs may object to this proposal because, in theory, defendants could submit false statements, make general denials, or withhold documents. As to false statements, requiring that they be under penalty of perjury, coupled with the fact that in many cases there will be multiple persons with knowledge of the truth, should minimize, but probably not completely eliminate, that possibility. As for general

---

[91] Rules Enabling Act of 1934, Pub. L. No. 73-415, 48 Stat. 1064 (codified as amended at 28 U.S.C. § 2072 (2006)); Rules of Decision Act, 28 U.S.C. § 1652 (2006).

denials and withholding documents, courts will have to examine those issues on a case-by-case basis to determine whether the defendant acted in good faith and was entitled to the protections afforded by this proposal for doing so. Put another way, this proposal is not perfect, but it is better for plaintiffs than the most likely reading of *Iqbal*, and it is a reasonable tradeoff that responds to whatever legitimate objections defendants actually have.