

**THE NECESSITY OF TRADEOFFS IN A PROPERLY FUNCTIONING
CIVIL PROCEDURE SYSTEM**

Alan B. Morrison¹

When judges, lawyers, and law professors discuss tradeoffs, it is usually in the context of debates about substantive policies. 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, and 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 made visible and their bases revealed, along with the reasons why one choice rather than another was made.

¹ Lerner Family Associate Dean for Public Interest & Public Service, George Washington University Law School. This essay is dedicated to the first year law students who struggle to understand the rules of civil procedure that sometimes seem to lack any coherence.

Rule 1 of the Federal Rules 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.” The problem is that just results often come slowly and/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 re-cast 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), where 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 issues at stake in the action, and the importance of discovery in resolving those issues.”

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, and unless that purpose or problem can be divined, the meaning and operation of the rule and any exceptions to it cannot be understood. But one purpose can rarely be advanced without some other purpose being set back, or reduced in significance, which 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 at least not something on which they generally 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 way to master a rule and learn how it should be applied. Put another way, even the most vanilla sounding rules are not “neutral” in that 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 sixty feet six inches from home plate, instead of exactly 60 feet.

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 rather than the other) and uncertain in advance. 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 rule or statutes, rather than developed on a case by case basis, 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 apply for 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 better to have a single set of Rules for all areas of the law because litigation will be simpler that way and 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, tradeoffs are not fixed, but are re-cast as circumstances change. As demonstrated below, discovery may be the clearest example of how the rules have evolved as discovery has become much more significant over the decades and, most recently, because electronic recordkeeping 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) that imposes on each party the duty to make certain affirmative disclosures. This obligation, first instituted in 1993, was added in an effort to reduce costs and lessen delays, and in doing so adjusted the adversary system so that for the first time a party had 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 committees sought to find the proper balance.

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)(1) “when the interest of justice so requires.” Rule 15(a)(2); see also Rule 15(c)(1)(C) allowing relation back of an amended complaint to add a defendant in certain circumstances provided that the defendant “will not be prejudiced” thereby. Similarly, 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. 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 choice of a rule focusing 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.²

Subject Matter Jurisdiction – Access to the Federal Courts

Article III of the Constitution allows federal courts to decide cases or controversies in only a limited set of circumstances – itself a tradeoff --- of which the most important are those “arising under this Constitution, the laws of the United States, and Treaties made under their Authority” (referred to as “federal question

² The same types of choices apply in substantive areas where Congress has chosen to make the tax laws very specific and the antitrust laws much more general.

jurisdiction”) and controversies “between Citizens of different States” (“diversity of citizenship jurisdiction”). That same Article also provides that Congress shall determine the jurisdiction of the lower federal courts, and it has chosen not to grant those courts the full extent of the power available under Article III.

Thus, when Congress created the federal trial courts in 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 having federal, rather than only state courts, decide issues arising under federal law in the first instance. But it did not open the federal courts to all such, 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. Then in two steps, first in 1976 and then in 1980, Congress reversed itself and eliminated the amount in controversy requirement. 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 many federal question cases (subject to possible review in the Supreme Court) and because the burden on the courts of determining whether some federal constitutional claims were “worth” more than \$10,000 was not worth the benefit of keeping a few “small” cases out of the federal courts (**Fifth Avenue Parade – Priceless is not worthless**).

For diversity cases, Congress has continued its gatekeeper role by limiting federal court jurisdiction to larger cases, now those in excess of \$75,000. The generally accepted reason for diversity jurisdiction in the Constitution is a fear that state courts would unduly

favor local residents, 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.³ In theory, that would suggest that all diversity cases can be heard in federal court, but Congress has consistently rejected that conclusion, largely because of the limited number of federal judges who might be overwhelmed if every diversity case could be brought in federal courts. Presumably, local bias would be present even in cases involving below \$75,000, but 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 it, all parties on one side must be citizens of states different from all parties on the other side in order for the requirement of diversity to be satisfied. The statutes governing diversity jurisdiction do not mention complete diversity, but the Supreme Court interpreted them to require it,⁴ and Congress has never rejected that reading, although 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 vs. state court for diversity cases is found in the Interpleader Act of 1925, 28 U.S.C. § 1335, under which the federal courts have jurisdiction so long as there is diversity between any two adverse parties. The Act was passed to deal with situations in which there are several

³ A similar rationale also explains why the Constitution gives aliens access to federal courts, but it does explain why Congress has allowed a U.S. citizen 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, even though there is no arguable out-of-state prejudice to the plaintiff if the case were handled in her own state court.

⁴ *Strawbridge v. Curtiss*, 7 U.S. (3 Cranch) 267 (1806).

claimants to a fixed fund, often an insurance policy, and the claimants live in different states, such that it would 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 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 relaxed the rule and allowed minimal diversity and set the amount in controversy at \$500. That amount has never been raised even though Congress has raised the general amount in controversy requirement in general diversity cases several times. 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 and/or claims. The Supreme Court had decided a series of cases, generally finding that these additional parties could not be added,⁵ although it was more generous with additional claims against existing parties.⁶ The result was that some parties were forced to litigate the same basic dispute in both federal and state courts⁷, which produced additional costs and some procedural unfairness. It was not disputed that, because there was subject matter jurisdiction for at least one claim in each case, there was no

⁵ *Aldinger v. Howard*, 427 U.S. 1 (1976).

⁶ *United Mine Workers of America v. Gibbs*, 383 U.S. 715 (1966).

⁷ *Finley v. United States*, 490 U.S. 545 (1989).

constitutional barrier to relaxing the rules of complete diversity, and the issue was which tradeoffs were appropriate to make, recognizing that any relaxation of these requirements would add to the caseload 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 can be added if they are part of the same case or controversy. 28 U.S.C. § 1367(a). But for diversity cases, the next subsection specified certain Rules under which claims and parties could not be added in diversity cases, thereby maintaining the old balance, which meant that efficiency of litigation gave way to keeping cases (or at some parties and some claims) out of the federal courts. Unfortunately, Congress did a very poor job of drafting, and left out of the list of 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 exclude those provisions meant that the general rule applicable to non-diversity cases applied and so additional, non-qualifying parties could be joined under those rules. *Exxon Mobil Corp. v. Allapattah Services, Inc.*, 545 U.S. 546 (2005). As the dissent pointed out, looked at 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 class actions, but exclude much smaller cases affected by subsection 1367(b), which is what the majority said Congress had done. 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 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 the *Exxon* case, Congress passed the Class Action Fairness Act (CAFA)⁸ that greatly expanded the ability to file and remove diversity class actions in federal court. Prior Supreme Court rulings interpreted the diversity provision on jurisdiction to hold that all members of the class had to have the requisite statutory amount in controversy,⁹ which meant that very few class actions based on state law could be brought 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 21st century, it was defendants who were seeking refuge in federal courts from class actions filed in state courts. Defendants generally thought that federal courts were less likely to grant class certification than were at least some state courts (mainly those where plaintiffs elected to file). Defendants also wanted to be able to obtain consolidation of 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. In addition, the federal removal statute also limited removal in diversity cases to those cases in which none of the defendants was a resident of the state in which the case was filed. 28 U.S.C. § 1441(b). That exception 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. See 28 U.S.C. § 1332(d). As to amount in controversy, it became \$ 5 million total, provided that there

⁸ Pub. L. 109-2, 119 Stat. 4 (2005).

⁹ *Zahn v. International Paper Co.*, 414 U.S. 291 (1973).

were 100 or more class members. As for diversity, CAFA opts 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 and the citizenship of the defendants is irrelevant. 28 U.S.C. § 1453. 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, 28 U.S.C. § 1332(d)(2) & (3), but the extent of the ability to affect remands is seriously in doubt, and in all likelihood, those provisions will prove to be relatively unimportant.¹⁰

Venue & 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 venues, statutes or in some cases rules provide some limits on where the case may be brought. Section 1391 of Title 28 imposes some limits on the choice of venue in the federal courts, but those limits are quite relaxed. Congress has also concluded that, even though venue is supposed to be a proxy for convenience – notice the similarity of roots for the two words – there are some cases in which the forum chosen meets the venue requirements and the court has personal jurisdiction over the defendant, but in the words of 28 U.S.C. § 1404(a) the case should be transferred to another forum where the case might have been brought “[f]or the

¹⁰ 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 75 natural persons who died (but not just injured, although injured persons may intervene) in a single discrete accident in certain prescribed situations.

convenience of parties and witnesses, in the interest of justice.” 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 on his chosen forum. Despite these obstacles, cases are transferred because Congress has made the judgment that, in some 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 pre-trial proceedings. Under 28 U.S.C. § 1407, Congress created the Judicial Panel on Multidistrict Litigation that has the power to transfer cases involving common questions of fact in the federal system to a single judge, but the location chosen need not be convenient for parties and witnesses – and indeed is often not. The rationale for this provision is to consolidate pre-trial matters, ranging from discovery to class certification to motions for summary judgment, before one federal judge to assure consistency and increase efficiency. The interests of individual plaintiffs (and perhaps even some defendants) in handling these pre-trial matters in a more convenient forum are submerged to the efficiency interests of the courts and all the parties with similar claims. In *Lexecon Inc. v. Milberg, Weiss, Bershad, Hynes, & Lerach*, 523 U.S. 26 (1998), the Court ruled that section 1407 required a re-transfer back to the original forum for trial, and while Congress has debated changing that rule, to date it has not acted, meaning that the trial judge who is most familiar with the case cannot try it.

Beginning the Lawsuit

The Complaint

The rule writers 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,” a practice commonly known as “notice pleading.” The Rule also requires that the complaint include a statement as to the basis for subject matter jurisdiction, but since federal courts are courts of limited jurisdiction, that is hardly a surprising or a 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 mandated. 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. And even though a complaint initiates a lawsuit, 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 be able to start a lawsuit in federal court. The continued viability of those tradeoffs in light of recent Supreme Court decisions interpreting Rule 8 is discussed below. 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 defendant drove his car negligently in striking plaintiff’s vehicle, without specifying that defendant drove too fast, ran a red light, or was talking on his cell phone. And 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, for example in a defective product case, where the defendant has exclusive access to all or at least the most important evidence on what caused plaintiff's injury, a general allegation of product defect 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 because he and his lawyers simply 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 even major issues in the case, even though in some cases defendant will have to hire a lawyer and spend time to rebut a case that has no basis. To many defendants, that tradeoff seems unfair, but the counter argument is that it is more unfair to permit a defendant to avoid liability because the plaintiff cannot learn enough to file a case where the defendant has the key evidence under its control.

There are a few situations in which there is a different tradeoff, in particular 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 is sometimes 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 rationale is hard to defend, and 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 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 competing needs of a defendant to be able to prepare its 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 believes that proof can be obtained through discovery. Rule 11. In short, not very much, which many defendants argue is an ill-advised tradeoff. Whatever one thinks of that tradeoff, there is no doubt that it helps plaintiffs at least get to court and stay in court until they can take discovery, a result that a higher level of required pre-filing 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 counter-productive. This example 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. 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 was generally 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 determining whether the Court's interpretation of Rule 8 was proper, there can be little doubt that the decisions in *Bell Atlantic v. Twombly*¹¹ and *Ashcraft v. Iqbal*¹² imposed a pleading standard for complaints that gave greater weight to the interests of defendants in avoiding discovery and in defending 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, re-set the balance in Rule 8 by re-interpreting 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.

¹¹ 550 U.S. 544 (2005).

¹² 129 S. Ct. 1937 (2009).

The Rules Committee has begun an examination, and in the meantime, Congress is considering whether to restore what it considers 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, as in effect a means of testing the waters. If the defendant did not provide a basis for contesting plaintiff's factual allegations, plaintiff's burden of pleading in the complaint would be minimal, but if defendant chose to disclose its version of the case, 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 are hopefully a means of increasing the fairness of the process to plaintiffs as they see the law after *Iqbal* and to defendants, as they saw it before *Twombly*.

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 can exercise personal jurisdiction over a defendant, there still remains the issue of by what means service of the complaint must be made. Again, assuming no constitutional problems, 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 defendant's residence sufficient? What about on one of defendant's children, or might the answer depend on the age of the child? What about others residing with 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 Rules required that someone in the office of the United States Marshall for the District in which the case was filed did all the personal service of process. 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 ___ civil filings in 1938, when the Rules went into effect, to ___, in** 1980 when the predecessor of current Rule (c)(2) was changed to allow any person over the age of 18, 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 Rule writers were willing to accept some additional litigation as a tradeoff for the reduction in cost of service of process and the delays that ensured when only the Marshall's Office was permitted to do the job. In addition, service by mail is acceptable in some circumstances now,¹³ not because the mails are perfect, but because the risk of non-delivery is

¹³ Rule 4(d) (1)(G)(waiver of service); Rule 4(e)(1) (when service by mail is allowed by state law); and

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 email) is the norm for similar reasons of cost and efficiency.

Similar kinds of 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. It is acceptable to leave the complaint with a person of suitable age at the home of defendant, Rule 4(2)((2)(B), because 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 plaintiff prevails. The presumption that allows service on others who live with the defendant has less force when the recipient is not a resident of the home, but is simply a co-worker, and so that tradeoff has not been included in the Rules. But when the complaint relates to the work of a business or government agency, other more 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 do so. These kinds of tradeoffs differ from those made in creating the rules governing the contents of a complaint because they do not make it easier for one side to remain in court and perhaps even 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.

Rule 4(i) (part of service on United States and its officers and agencies may be made by registered or certified mail).

Defendant's Response Options

Generally the defendant has twenty 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). The Rules could have given defendants a “reasonable time” to reply, which is what happens in the real world since motions for extension are routinely agreed to by 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 extensions are granted, the actual number of days is not crucial, but simply sets a default time period and gives 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 or statutes to support their claim. To the extent that may seem 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, which, in non-legalese, means that the plaintiff has no legal claim based on the facts set forth in the complaint. This kind of motion is entirely based on the absence of a viable legal theory for the plaintiff, and it requires defendant to accept as true all of the facts alleged in the complaint. Thus, in an auto accident case, where the claim is that the defendant, who was a 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 theory 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. Or, 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 passenger-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 can not respond 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 legal arguments of defendants, often with just a modest extension of time beyond the 10 days allowed in the rules. In some cases, the basis of the motion will be that the plaintiff failed to allege an essential element of the claim – such as that a federal antitrust violation restrained interstate commerce – in which case, 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 be given a chance to prove her case. If an amended complaint is filed, defendant can once again move to dismiss, and if that motion is granted, 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 without merit, but it at least gives them some opportunity to get rid of a case early in the process, before the expenses of discovery and, perhaps, a trial are incurred. And whether it is in fact adequate is of less significance than the fact that the tradeoff is one that is 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.

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. The reason 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, could make the risk so great that the case is not worth bringing. This pro-plaintiff rule also means that some less than fully meritorious cases will be brought and that defendants will have to pay their counsel to get them dismissed. The tradeoff is nonetheless considered on balance to be 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 the bringing of certain kinds of lawsuits – the most significant being claims for violation of civil rights – more attractive, and hence a different, one-way tradeoff on attorneys' fees was added to the mix. *See e.g.*, 42 U.S.C. § 1988.

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 fund to pay the lawyer. 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 the lawyer if he loses. Contingent fees are not permitted in criminal cases and in matrimonial matters, on the theory that the potential for misconduct in those kinds of cases creates improper incentives, thus making the tradeoff undesirable from a public policy perspective. 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 alter both of these rules on attorneys' fees (but not the exceptions), generally from the defense side, but there is little likelihood that the current tradeoffs will be altered in any significant respect in the United States in the near future.

Discovery

It may be difficult for civil litigators today to realize but, 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 a very limited discovery regime include the reality that discovery costs money and causes delays and 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 pre-trial discovery of what each side knows, which may also lead to earlier, or at least better-informed, settlements. But that choice is not without downsides, which include significant increases in costs and new (and sometimes excessive) burdens on the party that 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 pre-trial discovery, under which discovery is quite limited, **CITE**, subject to the constitutional requirement – itself a form of tradeoff – that the prosecutor furnish the defendant with exculpatory in his possession.¹⁴ Furthermore, other civil justice systems in other countries (and in most states for many years after the Federal Civil Rules were adopted, but not today) have made different tradeoffs and do not have anywhere near the kind of extensive discovery allowed under the 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 be 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, the Rule was changed to require the person on whom the request was served to answer it or 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 person, with a notice to take an oral deposition and a request that certain documents be brought to the deposition. The Rule has been changed so that a subpoena can seek only documents,

¹⁴ *Brady v. Maryland*, 375 U.S. 83 (1963).

without the necessity of having the custodian showing up and present them – another recognition that burdens and costs needed to be adjusted given the modern litigation practice model.

Several other aspects of discovery involve somewhat different tradeoffs. The principle of full discovery was tested early on in *Hickman v. Taylor*, 329 U.S. 495 (1947). 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. The Supreme Court ruled that the request was for a category of documents that it labeled “attorney work-product” and held that, because plaintiffs could interview or depose those witnesses, they should not be allowed, in effect, to freeload on the work of defendant’s counsel and, perhaps more importantly, gain 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. 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 and is now in Rule 26(b)(3). Like the original Court opinion, that Rule is not absolute, but allows discovery if there is a showing of substantial need and an undue hardship to obtain the information elsewhere, but 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, plaintiff could ask defendant the

name of any expert that defendant had consulted regarding the case, and then 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 now limit discovery to experts that a party has identified as a potential trial witness, and a deposition can be taken only after the witness has provided the opponent with a report concerning her proposed testimony. Rule 26(b)(4)(A). The expert witness is entitled to be paid at his or her commercial rates by the party taking the deposition, including for the time preparing for and attending the deposition. Rule 26(b)(4)(C). These current tradeoffs are fairly recent addition to the Rules, reflecting the increase use of experts.¹⁵

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." Rule 26(b)(1). 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, 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 gave was to require that information sought be relevant to the "claims or defenses" in the case, while allowing discovery as to the "subject

¹⁵ There is another change to Rule 26 that is being considered by the Supreme Court and on which action is expected by May 2010. Under it communications between counsel and their testifying experts would generally not be 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 regarding discovery of witnesses who may provide both factual and expert testimony are a further indication of the ever-shifting nature of the various tradeoffs that the rule on discovery from retained expert witnesses contains.

matter” only with court approval, for “good cause”. Rule 26(b)(1). 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, and added an explicit cost-benefit type analysis that allows the judge to decide when enough is enough. This re-balancing 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 recordkeeping has made caused the rulemakers to add provisions specially calling 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. Rules 16(b)(3)(iii); Rule 26(f)(3)(C), Rule 34(a)(1)(A) and (b)(2)(D)& (E), and Rule 37(e).

Joinder of Claims & Parties

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, to 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 those kinds of joinders 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 is not simply a defense to the main claim, but seeks affirmative recovery for the defendant. Rule 13(a) provides that, if a defendant has

a 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 plaintiff has 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 this controversy.

Similar kinds of 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 and/or efficiency (although those are not the terms actually used). If the requirements of the Rule are met, the party will be joined 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 section 1367(b) precludes the use of supplementary jurisdiction to bring in additional non-diverse parties under Rule 19, in effect, re-affirming 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 plaintiff nor 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 essay. What is important is that the Rule allows the interests of third parties and the systemic interest in overall efficiency to trump the interest of the existing parties in confining their lawsuit as they have framed it.

Class Actions

The 1938 version of the Rules had class actions, but it was not until 1966 that Rule 23 had real teeth in it. 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 the affirmative step to join the class, but having the right to exclude themselves from the class or “opt-out”. Given the inertia among class members, especially in cases with small claims, insisting on “opt-in” classes would mean that most cases would not go forward, which is plainly what defendants prefer. Because of the amount in controversy requirement for diversity cases, which, until CAFA was that each class member had to have the requisite statutory claim size, most classes were based on federal law where many statutes had no required amount in controversy, and where, after 1976, the \$10,000 general requirement was eliminated for all federal causes of actions. 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, and not just have laws on the books that do little to help their intended beneficiaries.

The tradeoff was tempered by certain burdens that were placed on the named plaintiffs in class actions or, more realistically, their lawyers. The first and most important was the requirement that, upon class certification, personal notice (which meant notice by mail) had to be provided to each class member who could be reasonably identified and that the class (counsel) had to pay for it. 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.¹⁶

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

¹⁶ *Eisen v. Carlisle & Jacquelin*, 417 U.S. 156 (1974).

Erie & State Law

The statutes granting jurisdiction over diversity cases do not establish what law should apply in those cases. By definition, there is no federal statute, treaty or constitutional provision on which the claim is based, but 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 100 years, until the Supreme Court ended it in *Erie Railroad v. Tompkins*, 304 U.S. 64 (1938), in which it held that state law, both common law 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, also promulgated by the Supreme Court, became effective, and it was clear from the outset that at least most of those Rules were procedural, not substantive, and hence would govern if there were a conflict between them and otherwise applicable state law.

The conflict was starkly raised in *Hanna v. Plumer*, 380 U.S. 460 (1965), in which the issue was whether the Massachusetts law that required that a complaint be served personally on the executor of an estate trumped the federal rule that allowed the summons and complaints 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 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 since the claim arose under state law, the executor argued that the service rules were also substantive and hence the state law, and 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. There is surely one way in which the Court has made a tradeoff that it could have, at least in theory, avoided. 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 when any Rule is even arguably applicable. The Court has recognized the importance of uniformity – *Hanna* is an example where uniformity was a significant reason to support the result there – 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.*, 518 U.S. 415 (1996), 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 the *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 interpreting these statutes and its 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.

Appeals

There are three examples in the field of federal court appeals that involve clear tradeoffs. The first is the final judgment rule, under which an appeal may only be taken from a final judgment, a term that is fairly strictly construed. That approach involves very significant tradeoffs, denying an immediate appeal when a defendant's motion to dismiss a case for failure to state a valid legal claim is rejected and thereby triggers 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 and/or may cause the plaintiff to have to litigate in both the trial and appeals courts at the same time. Because most cases settle (although with a different balance when 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 approach is not universally accepted. New York, for example, allows a wide range of interlocutory appeals, 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 and/or the trial 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.¹⁷ The most significant category of those cases involves rejections of claims by government officials of various kinds of immunity, often from liability from 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 relieve that person from having to defend the case at all. 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 as of right. For example, most denials of motions to dismiss or 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,

¹⁷ *Cohen v. Beneficial Indus. Loan Corp.*, 337 U.S. 541 (1949). Just this term, the Court unanimously rejected a claim that the denial of a claim of attorney-client privilege fell with the collateral order rule and denied the right to an interlocutory appeal. *Mohawk Industries v. Carpenter*, ___U.S. ___ (No. 08-678, decided December 8, 2009).

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 to hear it or not. 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 30 days from entry of the final judgment, 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 28 days from entry of judgment, and, unlike almost every other civil motion, that period (which was 10 days until December 1, 2009) cannot be extended. Rule 6(b)(2). Once the judgment becomes final, the 30 days for filing an appeal can be extended only for very limited and quite specific reasons, generally those for which the appellant is not responsible, such as the clerk not sending out the final order. Federal Rule of Appellate Procedure 4(a)(5). But even then, the forgiveness is limited in duration and the terms of the exception are quite confined. And most recently, the Court found an appeal untimely where the error had been made by the judge, and not objected to or even noted by the opposing party, even though the appellant had acted in reliance on the erroneous

additional extension given by the judge and so failed to file on time.¹⁸ In the Court’s words, 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 deadline for an appeal that is any more fixed than any other deadline, and Congress could, if it wished, change the statutes and/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 that the Court ascribes to those who wrote those statutes and rules.

Third, the degree of deference given to decisions of trial courts and juries is another example of a tradeoff on appeals. One tradeoff is made by the Seventh Amendment, which forbids the re-examination of facts found by a jury. As a result, there is a very high – some would say near-conclusive – tolerance for jury error as a lesser evil than having appellate judges substituting their views for that of the jury. But when the fact finder is the trial judge, the standard of review is the somewhat more rigorous “supported by substantial evidence” test under Rule 52(a)(6) for questions of fact, indicating less of a willingness to defer to a single judge’s view of the facts than to those a jury of six (or more). There are those who would argue that, despite the advantage that a trial 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

¹⁸ *Bowles v. Russell*, 127 S. Ct. 2360 (2007).

appellate panel were better able to reach the correct result, the same 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, a 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.

The added burden and the 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 to that gained from seeing witnesses who testified on factual issues. In fact, the trial judge often has less time to consider legal issues than do appeals 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 the issue is one of law that is not corrected on appeal, that may, as a practical matter, bind many others in similar situations. For each of these standards of review issues, different people might make different tradeoffs, that might produce different standards, but whichever way the rule ends up, there will inevitably be tradeoffs.

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 main goal of this essay 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.

Its secondary goal is to urge that those who write the rules, and courts that interpret them, should 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 then to apply the rule in practice. 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 one only reading, especially when a case is in a court of appeals, let alone the Supreme Court. By pointing to the part of the tradeoff supporting the outcome, courts will both 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 in civil cases in federal court, and those who support giving plaintiffs with a good faith belief in the merits of their case a chance to prove it in court, fear that the legacy of *Iqbal* will be to enable defendants who have exclusive possession of key evidence to obtain dismissal of the complaint because plaintiffs will not be able to supply the missing allegations that *Iqbal* seems to require. 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 seeming 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 the filing. But if the defendant does not choose to provide that evidence to plaintiff, the court, in ruling on a motion to dismiss, could not grant the motion when the missing factual matters were under the sole 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 discuss other aspects of the proposal. Before turning to those cases, a brief examination of a very common type of case in federal court, where the pleadings are always conclusory, and sometimes not plausible, yet are never dismissed under Rule 12(b)(6): actions under the Freedom of Information Act (FOIA), 5 U.S.C. §552.

A typical FOIA complaint alleges that the defendant agency has certain records that the plaintiff requested, that the defendant did not furnish the requested records (and in some cases did not give any factual and/or legal basis for its denial), and that the denial was without basis in law, *i.e.*, none of the enumerated exemptions properly applies. 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 complaint is (“unlawful” or “without basis in law” is pretty conclusory) or how implausible it is that DOD has improperly classified the records, let alone that a court would so find. The reason why those cases are allowed to go forward and why the courts insist that the defendants come forward with factual support for their claims 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 now pending before the 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 impact on African-American applicants and lacked a business justification, thereby violating Title VII. The issue before the Court is the timeliness of the claim, and in that case Chicago did disclose the adverse impacts when it announced the test results. But, as

is more likely, especially for private employers, there would have been no such announcement, and the plaintiff would have filed charges with the EEOC, the defendant would not have responded, and a right to sue letter was issued. After *Iqbal*, a plaintiff who alleged only that the cutoff score had a disparate impact on African-Americans and lacked a business justification might have the complaint dismissed for including merely “conclusory” allegations, even though all the detailed information was in control of the defendant. Under this proposal, if the same sequence was followed, the court would be forbidden from dismissing 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 on justification when it had the opportunity to do so before suit was filed.

Or assume that a Toyota owner, whose vehicle suddenly accelerated to 90 mph and crashed into a tree, seriously injuring the driver, alleged that the car was negligently designed and manufactured and that there were breaches of various warranties, but no specific statements as to the actual cause of the acceleration. Under at least some readings of *Iqbal*, the plaintiff could have the case dismissed for lack of specificity, beyond the kind curable after a Rule 12(e) motion. However, under this proposal, if the plaintiff gave Toyota a copy of the proposed complaint containing those allegations, and it did not respond, the plaintiff would be able to defeat a motion to dismiss and commence discovery (assuming that there was not some other legal basis on which the complaint might be dismissed).

In *Iqbal* itself the Court held that the allegations that the Attorney General and the Director of the FBI approved the allegedly discriminatory policies were too indefinite to

be allowed to go forward. The plaintiffs had made every reasonable effort to ascertain the facts as to what involvement, if any, those two officials had, including even in discovery in that case, but they were rebuffed. Of course, both defendants know (or at least their officials files would show) whether they had approved any policies regarding the detention of aliens after 9/11 of the kind set forth in the complaint; they simply chose not to provide that information. The case was allowed to go forward against the remaining defendants, but not against the Attorney General and FBI Director, and was settled for \$265,000.

If this proposal had been in effect, 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 consequences if the Government remains silent at the administrative level.¹⁹ 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 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 12(b)(6) motion that the dissenters and Iqbal's counsel agreed was available. But if the defendant provided relevant 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

¹⁹ 28 U.S.C. § 2675(a).

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, or submitted any other evidence that someone else had approved the policies, plaintiff would not be stuck with the unsupported denial because those statements would not have been subject to cross-examination and could be very general. In other words, where there is likely to be a paper or email 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 of an agreement not to be plausible, especially in light of other allegations that pointed toward conscious parallel conduct that is lawful. 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. But if the defendants each responded with affidavits from senior corporate officials, based on personal knowledge and an investigation described in the affidavits, that no meetings on this subject ever took place and that there are no records (paper or electronic) that support a conclusion that an agreement among the defendants exists, the plaintiff could still file the complaint, then, unless the plaintiff were able to be more specific than the plaintiff was in *Twombly* itself, the complaint could be dismissed under Rule 12(b)(6).

The advantages 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 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, with a pre-suit exchange in process, in which defendants would have a stake in participating, not stonewalling, settlement might be discussed, based on facts and not just suspicions.

Some of these benefits will also accrue to defendants, but they will 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 (engage with plaintiff) or not. Engaging means some discovery, although much less than in court because the process is voluntary and can be stopped at any point. But there is no reason why this limited discovery can not 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.

Ideally, there would be an automatic tolling of the statute of limitations while the exhaustion option was pending, but given the limits of the Rules Enacting Act and the Rules of Decision Act, it is doubtful that the rules themselves could accomplish that. Nothing would prevent the parties from entering a tolling agreement, and if the exhaustion period were only 60 days, that should not cause many statute of limitations problems, especially because it would ordinarily be used promptly after the injury was discovered, before the statute of limitations 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 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 that defendants actually have.