

Allocation of Discovery Costs and the Foundations of Modern Procedure

Martin H. Redish

The Federal Rules of Civil Procedure are rapidly approaching their 75th birthday, which will come in the year 2013. 75 years is a long time, and while the Rules have of course been amended significantly at various points over the years, their basic structure remains largely the same as in their original formulation. When first promulgated in 1938, the Rules had an immediate and dramatic impact on civil adjudication by replacing long accepted procedural practices with very different methods of resolving disputes. There can be little question that the new system, spearheaded by the genius of Advisory Committee Reporter Charles Clark,¹ radically altered not only the actual procedures themselves, but also the underlying set of values that had previously rationalized our procedural system.

In place of the draconian requirements of the demanding fact pleading standard, which required a plaintiff to know all of the circumstances surrounding his injury in detail at the time of the pleading, the new Federal Rules demanded considerably less at the pleading stage. The information that was unavailable at the pleading stage could now

¹ See generally Michael E. Smith, Judge Charles E. Clark and the Federal Rules of Civil Procedure, 85 Yale L. J. 914 (1976).

be gathered through a complex system of court enforced discovery.² This dramatic change was premised in some sense on the notion that “[m]utual knowledge of all the relevant facts gathered by both parties is essential to proper litigation.”³

Despite the admiration that those who shaped the Federal Rules no doubt deserve, serious problems have remained. For the most part, those problems have revolved around the significant burdens that the Rules’ elaborate discovery process has imposed on litigants and the judicial system. In many ways, discovery is reminiscent of the invention of fire. Like fire, if used properly the discovery system can dramatically improve our situation. If abused, however, discovery can cause immeasurable damage. This point is underscored by the Supreme Court’s lingering concern over the serious burdens caused by the elaborate discovery process that represented the original Federal Rules’ most significant innovation. Designed to enable litigants to gather the information necessary to facilitate accurate decision-making and the effective vindication of substantive rights, the discovery process has a dark side that seems to have been largely undervalued at the time of the Rules’ framing. At least in an important category of litigation—those cases in which significant amounts of discovery are likely to take place—the costs and burdens inherent in the discovery process threaten to give rise both to serious inefficiencies in the adjudicatory process and to a potentially pathological and coercive skewing of the applicable substantive law being enforced.

² See *Hickman v. Taylor*, 329 U.S. 495, 500-501 (1947).

³ *Id.* at 507.

The Court clearly reasoned in its recent pleading decisions that unless the pleading standards effectively perform some form of meaningful gate-keeping function, the harms caused by excessive and burdensome discovery could easily overwhelm the adjudication in much of modern high stakes litigation.⁴ Yet even with the pleadings standard performing this filtering function, the fact remains that in cases that are allowed

⁴ See *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544, 559 (2007):

It is no answer to say that a claim just shy of a plausible entitlement to relief can, if groundless, be weeded out early in the discovery process through “careful case management”... given the common lament that the success of judicial supervision in checking discovery abuse has been on the modest side. See, e.g., Easterbrook, *Discovery as Abuse*, 69 B.U.L.Rev. 635, 638 (1989) (“Judges can do little about impositional discovery when parties control the legal claims to be presented and conduct the discovery themselves”). And it is self-evident that the problem of discovery abuse cannot be solved by “careful scrutiny of evidence at the summary judgment stage,” much less “lucid instructions to juries,” *post*, at 1975; the threat of discovery expense will push cost-conscious defendants to settle even anemic cases before reaching those proceedings. Probably, then, it is only by taking care to require allegations that reach the level suggesting conspiracy that we can hope to avoid the potentially enormous expense of discovery in cases with no “‘reasonably founded hope that the [discovery] process will reveal relevant evidence’ ” to support a [Sherman Act] § 1 claim.

to proceed beyond the pleading stage, the burdens and costs of discovery are likely to continue to be substantial. The problems of excessive discovery, then, remain a significant concern.

Revisors of the Federal Rules have, over the years, sought to remedy the harms of discovery. As the Supreme Court itself has recognized, however, at least in the “mega-litigation” context, serious problems remain.⁵ Yet there do exist hitherto untapped alternatives. It is my position that foundational precepts of economic, moral and political theory dictate a dramatic alteration in the structural operation of the discovery process. If implemented, this change would undoubtedly reduce the costs and burdens of the process while preserving the bulk of its beneficial functions. This alteration would recognize that the costs of discovery are, from the outset, properly attributed to the requesting party, rather than the responding party. Indeed, classic notions of quantum meruit—long recognized as an indisputable moral and legal dictate in the law of contracts—permit no other conclusion. Were this alteration in the nature of the discovery process to be implemented, an immediate economic externality—one that currently plagues all discovery requests—would immediately be removed. As a result, the discovery system would be relieved of most forms of even non-abusive “excessive” discovery requests—discovery that is simply not justified on the basis of a rational cost-benefit analysis.

Discovery and the Substantive-Procedural Balance

⁵ Id. at 558-559.

It was quite clear, when the Federal Rules were first promulgated, that by simultaneously reducing the barriers imposed by the fact pleading requirement and establishing a complex set of court-enforced information gathering devices, Judge Clark and the Advisory Committee were attempting both to increase procedural fairness to plaintiffs and employ procedure as a more effective means of implementing substantive law. Because of the potential burdens of discovery, however, an unduly lax pleading standard cannot be imposed without seriously skewing the substantive-procedural balance towards pathological over-enforcement of the substantive law. To be sure, discovery is designed to play—and does play—an important role in the effective procedural implementation of substantive law by providing deserving plaintiffs with a viable means of acquiring the information necessary to enforce their rights through resort to the adjudicatory process. Nevertheless, the very same danger of pathological over-enforcement exists with regard to discovery.

The Problem of Discovery Costs

If left wholly unregulated, discovery can give rise to numerous procedural and substantive pathologies. In its most extreme form, intentionally abusive discovery effectively transforms the adjudicatory system into a means of facilitating legalized blackmail and extortion. The very threat of costly discovery likely induces rationally self-interested defendants to settle even non-meritorious suits for an amount smaller than the projected costs of discovery. Requests for wholly unnecessary discovery could thus

function in an extortionate manner, financially coercing a defendant into settling unjustified claims. It is fundamentally unfair to a defendant to allow the adjudicatory system to be employed against him as a weapon of extortion.

No one, presumably, would openly sanction discovery that is unambiguously abusive. The problem, of course, is to find ways to control such pathological discovery without either effectively destroying the beneficial effects of the discovery process or establishing control methods that are as economically inefficient as the abusive discovery itself. This has proven to be far more difficult than one might have hoped. It is to consideration of this difficult issue that my analysis turns.

The “Cost Allocation” Model and the Control of “Excessive” Discovery

Although the revisors of the Federal Rules have, over the years, diligently sought to control the abuses of the discovery process, few would argue that they have been successful in that effort. Puzzlingly, despite all of these efforts the revisors have never even considered, much less adopted, the system that would remedy the problems of discovery most effectively. Such a system would, simply, allocate discovery costs to the party who both initiates the discovery and benefits most as a result of it.

From the outset, it is important to understand that what is advocated here is not a process of cost shifting. Indeed, the very use of that word necessarily concedes that the inertia of cost allocation appropriately belongs on the responding party, and must be

shifted in order to have discovery costs attributed to the requesting party. Yet at no point has anyone—including those who drafted the Federal Rules in the first place—even attempted to rationalize the respondent-centric model of cost allocation that has dominated federal court practice since the Rules’ original promulgation. Were one actually to consider the issue afresh, it would be difficult to understand the assumptions inherent in a respondent-based allocation model. It is true, of course, that in the crudest, most concrete sense the cost is immediately incurred by the responding party, not the requesting party. But that fact, standing alone, in no way necessarily implies that even at that point is the cost appropriately to be attributed to the responding party.

One may best understand the point by consideration of a simple analogy. Assume a worker asks his co-worker to do him a favor and pick up lunch for him. The co-worker does so, paying the \$15 that the lunch costs. He then brings the lunch to the requesting worker; unless he was raised by wolves, the requesting worker will immediately thank the co-worker and reimburse him for his \$15 expenditure on his behalf. Is such reimbursement appropriately characterized as “cost shifting” in anything but the most concrete, technical and immediate sense? At any point in this hypothetical transaction, was the cost of that lunch appropriately viewed, morally or conceptually, as the co-worker’s cost, rather than the requesting worker’s cost? Long established principles of quantum meruit readily answer that question in the negative.⁶ The co-worker performed work on behalf of another, who was aware both that the co-worker was performing the

⁶ Martin H. Redish and Colleen McNamara, Back to the Future: Discovery Cost Allocation and Modern Procedural Theory, 79 Geo. Wash. L. Rev. 14-18 (forthcoming 2011), available at <http://ssrn.com/abstract=1621944>.

work on his behalf and, as a result, incurring costs on his behalf. The law of quasi-contract unambiguously dictates that in such a situation the cost is deemed that of the party on behalf of whom the work was done, not of the party who performed the work.⁷

In fundamental ways, the discovery process is identical to this hypothetical situation. The only differences are that in the case of discovery, the performing party is usually performing the work not out of the goodness of his heart but rather due to the coercive threat of court sanction if he fails to do so. Moreover, the work performed by the responding party will not only help the requesting party but often actually harm the interests of the responding party himself. These differences, however, make even more bizarre the seemingly universal but wholly unsupported assumption that discovery costs are appropriately attributed to the responding party, rather than to the requesting party.

It should be clear that as both a legal and moral matter, the costs of discovery are properly attributable, in the first instance, to the requesting party. By imposing the costs of discovery on the responding party, then, our system has effectively required the responding party to provide a subsidy to the requesting party. To be sure, assuming no constitutional problems,⁸ the system may choose to order such a subsidy. But because those who created the system implicitly—and inaccurately—assumed that the costs of discovery was properly seen as a cost to be borne by the responding party, our system has provided for a hidden subsidy, recognized by no one. At the very least, democracy demands that the decisions of those who make fundamental choices of social policy make

⁷ Id.

clear what those choices actually are, so a transparent debate of whether it is fair to impose such a subsidy may finally take place. This has never been done in the case of discovery costs.

Cost Allocation and the Externalities of the Discovery Process

Wholly apart from this complete lack of transparency, the implicit assumption that the costs of discovery are to be attributed to the responding party makes little sense, from any theoretical or practical perspective. This is particularly true when coupled with the broad scope of discovery in the age of informational technology. In addition to its moral and legal bases, attribution of the costs of discovery to the discovering party, rather than the responding party, is likely to have significant instrumental benefits, because it would cure what has long been a fundamental economic pathology plaguing the discovery process: the externality inherent in the choice to invoke discovery. Simply put, under the prevailing practice the cost-benefit decision whether or not to invoke the discovery process is made by a party who risks incurring no cost, only benefit, even though it is quite conceivable that the choice will impose a significant cost on others.

This lack of economic disincentive underscores what may well be a far greater harm to the system than intentionally abusive discovery: what can be most appropriately labeled “excessive” discovery. This concept includes discovery which, while not consciously interposed for purposes of delay or harassment, nevertheless gives rise to costs greater than its benefits in finding truth. It is widely recognized that the value of

finding truth cannot be considered in a vacuum, wholly divorced from the costs to which the effort gives rise.⁹ A rough judgment must always be made by some decision maker whether the likely benefit to come from the effort justifies the effort's costs. Yet when the responding party, rather than the requesting party, bears the costs of the process, the requesting party has absolutely no economic disincentive not to make the request, regardless of its costs. Indeed, given that it is the requesting party's opponent who will bear that cost, one might even suggest that in a perverse sense, the higher the cost the greater the incentive to invoke the discovery process.

This focus on the subtle but important differences between "abusive" and "excessive" discovery underscores the manner in which a reversal in the ex ante presumption of discovery cost attribution can function in a symbiotic manner with both the direct and prophylactic methods of discovery control. While those more judicially driven practices are more likely to punish or deter *abusive* discovery, the self-executing shift in discovery cost allocation is far more likely to deter the practice of *excessive* discovery.

The key social problem to which imposition of discovery costs on the requesting party might give rise derives from its inherently regressive nature: the poor will be more immediately and seriously impacted by such costs than will the rich. To be sure, this is also true of all litigation costs, though this fact has never caused us to shift all of the poor's litigation costs to the wealthier party. Moreover, particularly in the case of

⁹ See *Mathews v. Eldridge*, 424 U.S. 319 (1976) (adopting a test that balances systemic costs against goal of accuracy in determining procedural due process).

complex class action lawsuits, the real party in interest will not be the individual plaintiff but rather the plaintiff's attorneys, for whom the funding of such suits is simply a cost of doing business. In these cases, it would be wrong to see this alteration in discovery cost allocation as an inherently regressive practice. In any event, if there are particular substantive rights which the governmental body decides require procedural subsidization, that body may say so at the time it creates those rights. Therefore even if one were to find the regressive impact of this reversal in cost allocation to be a matter of concern, a wholesale rejection of the cost allocation model would not be justified.

Even if society were to decide to subsidize a poorer litigant's discovery in particular suits, it hardly makes sense to impose that cost on his opponent, rather than on society as a whole. Indeed, to allow a private individual's unilateral filing of a lawsuit to justify imposition of discovery costs on the defendant gives rise to serious constitutional concerns of due process. The Supreme Court has long held that due process prohibits the deprivation of a defendant's property absent meaningful judicial involvement in the determination of that defendant's culpability.¹⁰

A conceivable objection to the reversal of the current cost allocation model might be that such a practice would simply shift the externality, for under the new model the responding party will have no incentive to keep costs down. But it is the discovering party who sets the contours of the response by the scope of its inquiries or production requests. In an important sense, then, the outer limits of the costs that the responding party will incur are set out by the requesting party. In any event, there always exists the

¹⁰ See, e.g., *Fuentes v. Shevin*, 407 U.S. 67 (1972).

possibility of judicial intervention to determine that the submitted costs are excessive.

While it might be responded that such intervention would significantly increase the systemic burdens of the discovery process, it is highly unlikely that judicial intervention would be required in many instances. If the responding party knows that any excessive costs it incurs may well not be reimbursed, it is unlikely to risk incurring them in the first place.

Proposing Amendments to the Discovery Rules

While formal amendment of the pleading rules may not be essential, the same is not true of the discovery process. As explained earlier, the key to taming the discovery process is to understand that, in the first instance, the costs of discovery are appropriately seen as costs attributable to the requesting party, rather than the responding party. While in its current form Rule 26(c), authorizing the issuance of protective orders, is framed in a manner that vests broad discretion in the district court's hands to "shift" costs, such a power is only rarely employed. In any event, the point of the amendment would not be merely to *authorize* the court to *shift* costs, but rather expressly to attribute the costs, in the first instance, to the requesting party. Rule 26 should therefore be amended to state unambiguously that discovery costs are attributable to the requesting party, unless

applicable substantive law provides to the contrary or the court finds that a compelling reason for shifting the costs to the responding party exists.¹¹

Conclusion

First year law students have long been taught that law is not simple; there are invariably conceptual and practical complexities that must be carefully balanced. Though it is perhaps difficult for us now to see it, the genius of the drafters of the original Federal Rules was their ability to recognize those complexities and to seek carefully to balance the competing needs as a means of achieving a solution that takes all of those complexities into account.

¹¹ Beyond this amendment, it would also make sense for the Advisory Committee to consider possible alternative methods of directly controlling discovery. One such method that has been suggested is restriction of the scope of available discovery. For example, respected organizations have suggested that: “Discovery in general and document discovery in particular should be limited to documents or information that would enable a party to prove or disprove a claim or defense or enable a party to impeach a witness.” See, e.g., Am. College of Trial Lawyers & Inst. For the Advancement of the Am. Legal Syst., Final Report 7 (2009), Final Report on the Joint Project of the American College of Trial Lawyers Task Force on Discovery and the Institute for the Advancement of the American Legal System 14 (2009) at 8. That question, however, is an issue on which this Article takes no position.

Discovery Costs

The drafters of the Rules, of course, were only human, and humans make mistakes—especially in the process of revolutionizing an entire system. In the discovery process, their first mistake was their failure even to consider the question of to whom discovery costs were to be appropriately attributed in the first instance. Their second mistake was their flawed implicit assumption that the costs were properly to be attributed not to the party who is best able to economically internalize the costs and benefits of discovery, but to the party who has little or no control over those decisions. It is now time to correct their errors—and get ready to wish them a happy birthday.