PERFECTING FEDERAL CIVIL RULES: A PROPOSAL FOR RESTRICTED FIELD EXPERIMENTS

Laurens Walker*

I

INTRODUCTION

The Federal Rules of Civil Procedure are now more than fifty years old, and surely the time has come to review the process which governs their development. The fiftieth anniversary provides a strong symbolic reason for this review, but that is not all. The original task for rulemakers was to produce a whole set of civil rules to meet immediate basic needs, and the anniversary certainly suggests that this work was accomplished. Today, the salient task for rulemakers is to improve this system and to adjust it in response to new conditions. Toward this end, the present Rules can serve as a base line measure for comparison with proposed changes. Now that the work of creating a whole new system is complete, there is ample time to examine individual proposed changes.

It is the thesis of this article that the process which guided the development of the Rules through their first fifty years is not appropriate for the work that lies ahead. The chief deficiency is the lack of a systematic official plan to collect valid information about the likely impact of changes to the Rules before the adoption of general amendments to the Rules. A program of restricted field experiments should be adopted to predict the impact of proposed changes to the Federal Rules of Civil Procedure.

The conditions that Professor Charles Alan Wright observed in 1967 are little changed: “We know very little about how present procedures work. We know even less well what changes might produce improvement in the

Copyright 1988 by Law and Contemporary Problems

* T. Munford Boyd Professor of Law, University of Virginia. The assistance of Craig Burns, Preston Burton, Sally Groseclose, Robert Hupe, and Page Wittkamp in the preparation of this article is gratefully acknowledged. The helpful comments and suggestions of the Law and Contemporary Problems Conference participants are also acknowledged, with additional thanks to Michael Saks, who made suggestions at an early stage of the work, and to Maurice Rosenberg and Allan Lind, who provided suggestions after the Conference. The assistance of my colleagues Saul Levmore, John Monahan, and Glen Robinson is also acknowledged, with appreciation.


future.” Indeed, Judge Richard Posner recently proposed, as one criterion for evaluating procedural reform, that “[t]he success or failure of the procedure must be verifiable by accepted methods of (social) scientific hypothesis testing.” He explained, “I am unconvinced by anecdotes, glowing testimonials, confident assertions, appeals to intuition. Lawyers, including judges and law professors, have been lazy about subjecting their hunches—which in honesty we should admit are little better than prejudices—to systematic empirical testing.” Today, given the task of perfecting a mature system of procedure and given the parallel maturity of the social sciences, the legal system should no longer tolerate this lack of empirical testing.

Part II of this article reviews past research about the Federal Rules of Civil Procedure and arrives at three conclusions: (1) there has been very little empirical research on the Federal Rules; (2) most research has been done after change was implemented; and (3) the research methods most commonly used are not well suited to predict the impact of change. Part III briefly recalls the significance of the Federal Rules of Civil Procedure, chiefly by suggesting the effect of the Rules on the American public. Part IV examines the benefits and costs of collecting information under current practices. Then, Part IV examines the benefits and costs of using a system of national experimentation. Finally, this section of the article considers a system of geographically restricted field experiments and compares such a system with both the current practice and national experimentation.

Part V considers the potential argument that random assignment of cases to civil rules test conditions would be a denial of equal protection of the law, as guaranteed by the Constitution. Part VI discusses ethical issues relating to the proposed research in light of a recent Advisory Committee report that suggested criteria for carrying out research in the federal courts. Next, Part VII considers how to implement a system of restricted field experiments. This section of the article concludes that an abandoned proposal to permit temporary research projects under local civil rules should be revived and adopted, provided that the proposal is clarified to permit random assignment of cases to test conditions. The conclusion urges rulemakers to implement the proposed changes, arguing that a system of restricted field experiments would serve the public interest.

II

PAST EMPIRICAL RESEARCH ABOUT FEDERAL CIVIL RULES

A comprehensive search for empirical research which focuses directly on one or more of the Federal Rules of Civil Procedure yields several striking conclusions.6 First, very little empirical study of the Federal Rules of Civil Procedure has been conducted. During the period of their development and the fifty years since the adoption of the Federal Rules, only about thirty different studies have been reported.7 An average of approximately one study

6. The search covered the years 1954 to 1987 and was focused on research concerning specific Federal Rules of Civil Procedure and including a substantial empirical component. These limits, adopted because of the topic of this article, excluded a number of notable projects which had a different focus. For example, the pioneering pretrial conference study of Maurice Rosenberg was not included because the Rosenberg study investigated a proposed New Jersey pretrial conference rule, although certainly the results of the study had important implications for federal procedure. See M. ROSENBERG, THE PRETRIAL CONFERENCE AND EFFECTIVE JUSTICE (1964). For similar reasons, the search did not include the various major studies of the Wisconsin Civil Litigation Research Project, as reported in D. TRUBER, J. GROSSMAN, W. FELSTINER, H. KRITZER & A. SARAT, CIVIL LITIGATION RESEARCH PROJECT: FINAL REPORT (1983). Also, the now considerable body of procedural justice research is not included. See, e.g., J. THIBAUT & L. WALKER, PROCEDURAL JUSTICE: A PSYCHOLOGICAL ANALYSIS (1975); E. LIND & T. TYLER, THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE (1988). The criteria also excluded important evaluations of court-annexed arbitration, as reported in E. LIND & J. SHAPARD, EVALUATION OF COURT-ANNEXED ARBITRATION IN THREE FEDERAL DISTRICT COURTS (1981). Also, important empirical studies about juries such as H. KALVEN & H. ZEISEL, THE AMERICAN JURY (1971), and M. SARK, JURY VERDICTS (1977), were not within the criteria.

every other year is scant empirical attention to the Rules. There are several plausible reasons for this small body of work. Empirical research is costly and time-consuming. Also, it remains relatively novel to legal academics. Finally, rulemakers have not often demanded empirical evidence about the likely impact of change, and have thus discouraged investments of money and time.

A second conclusion concerns the timing of research projects. Most of the empirical research about federal civil rules has been done after the general adoption of the rule or rules studied. Obviously, the public may have been, and may still be, subjected to poor quality rules for long periods. There is little reliable information about the impact of the current Rules. Also, this practice of adopting changes before research undoubtedly contributes to the research method problems discussed next. General implementation of rule changes before testing causes great difficulty in determining the impact of change.

The final conclusion involves research method. The survey is the method most often used to investigate the Federal Rules of Civil Procedure. Some surveys have collected information from judges and attorneys. Others have been archival studies which have relied chiefly on court records. Still other surveys have used both types of sources. Although this technique can certainly produce valuable information, the survey presents severe problems for determining cause and effect relationships. Yet, cause and effect is a key issue in efforts to improve legal procedures by changing rules. One study located was a "quasi-experiment," a research method which presents fewer problems for causal determinations. True experimental method, which provides a strong basis for causal judgments, was used in a study reported elsewhere in this issue, and in another recent project. However, no other true experiments about the Rules were found. Thus, most empirical research concerning the Federal Rules of Civil Procedure has used methods that prevent confident conclusions about the impact of change. Discerning the effect of Rule amendments has been left mostly to guesswork.

---

9. One pre-rule adoption study found in the search was Sand & Reiss, supra note 7. Another is reported in W. Glaser, supra note 7. For a third reported in this issue, see Rowe & Vidmar, supra note 7.
10. The characteristics of the methods briefly discussed here are examined in more detail in the text and notes beginning at p. 106.
11. See, e.g., W. Glaser, supra note 7; Sand, supra note 7; Shapiro, supra note 7.
12. See, e.g., Flanders, supra note 7; Nelken, supra note 7; Vairo, supra note 7. See also J. Simon, Basic Research Methods in Social Science 193 (1978) ("A survey need not be a survey of people. You can survey either people or things.").
13. See, e.g., P. Connolly, E. Holleman & M. Kuhlman, supra note 7; Ellington, supra note 7; Speck, supra note 7.
15. Rowe & Vidmar, supra note 7.
17. In 1960, Geoffrey Hazard undertook a general review of research about civil procedure. See G. Hazard, Research in Civil Procedure (1963). The scope of his work was not limited to federal
III

THE PUBLIC INTEREST RECALLED

Given the concerns about how procedures work and what changes might produce future improvement, it is appropriate to recall briefly the public interest in the quality of the Federal Rules of Civil Procedure. While the specific impact on the public cannot be measured, a simple consideration of the number of litigants required to follow the rules suggests that the impact is significant. During 1987, the latest year for which data are available, 238,982 civil cases were commenced in the U.S. district courts. Under the conservative assumption that the interests of at least two persons are affected by each case, during 1987 almost 500,000 persons were affected to some degree by the content of the Rules. Furthermore, recent trends increasing the complexity of litigation suggest that some of the suits filed each year in federal court may each directly affect the interests of thousands or perhaps millions of persons. But the significance of the Federal Rules (and changes to the Rules) is certainly not limited to federal court litigation, because the Rules have become a model for many states. According to a 1986 article, at least twenty-two states and the District of Columbia have civil rules that can be classified as "replicas" of the Federal Rules. During 1986, the most recent year for which data are available, 2,228,352 civil cases were filed in the courts of general jurisdiction of those twenty-three jurisdictions. Given the previous assumption, this amount suggests that in these jurisdictions in 1986, the replica rules affected the interest of more than 4,000,000 persons. The same study of Rule adoptions indicates that to a lesser degree, a number of other state systems of civil procedure are influenced by the content of the Federal Rules, and it is easy to see that a large number of state civil litigants are affected by the content of the Federal Rules. Surely, the interests of millions of Americans are involved in Federal Rule changes.

---

18. See supra note 3 and accompanying text.
23. Oakley & Coon, supra note 21, at 1369.
Having established the general problem and its significance, one must, as a vehicle for systematically comparing solutions, consider the benefits and costs of the present ways of evaluating Federal Rule changes and compare those practices with alternatives. The benefits of the various research methods will be described according to the probable validity of the cause and effect information that each method can yield. The concept of cost will include both the cost of imposing a bad rule on the public (error cost) and the cost of carrying out the research (transaction cost).

A. Current Practices

According to Cook and Campbell, the validity of causal research designs is commonly assessed by considering two subcategories, internal validity and external validity. The concept of internal validity describes the power of a research design to eliminate unintended causes of research results. In other words, considering the general topic of civil procedure, the internal validity of a study would typically indicate the power of a design to rule out causes other than rule changes when important effects are measured. Present research practices produce results of dubious internal validity, mainly because most of the research to date has been conducted through the use of surveys. According to Simon, the major disadvantage of the survey in cause and effect research is the "lack of manipulation of the independent variable." In procedure research there has typically been no controlled introduction of a rule change, and hence it is always possible that observed effects may be the result of some action or event other than the rule change. Of course, as Simon points out, it would be wrong to state that surveys never show causation. Nevertheless, the internal validity problem is so severe that conclusions about cause and effect—often so important in planning procedural changes—cannot confidently be based on such work.

24. Purely descriptive research can be very useful, but rulemakers concerned with improving civil courts must focus chiefly on questions involving cause and effect. Rulemakers typically need to know whether an amendment will produce some important impact such as a reduction in attorney fees, increased party satisfaction, more settlements, and so forth. Questions about impact necessarily implicate cause and effect. See Diamond, Methods for the Empirical Study of Law, in LAW AND THE SOCIAL SCIENCES 637, 665-66 (L. Lipson & S. Wheeler eds. 1986).

25. This term is a very loose adaptation of "the cost of erroneous judicial decisions" discussed in R. POSNER, ECONOMIC ANALYSIS OF LAW 517-21 (3d ed. 1986).


28. T. COOK & D. CAMPBELL, supra note 27, at 70-73; see also J. MONAHAN & L. WALKER, supra note 27, at 50-52.

29. J. SIMON, supra note 12, at 192 (emphasis in original).

30. Id.
In addition to surveys, one quasi-experiment and two true experiments have focused on the Rules. Quasi-experimental designs provide more confidence with respect to causal information. A comparison group of subjects is used to provide stronger internal validity, but because subjects are not randomly assigned to that group, serious threats to validity remain. The true experiment is the most powerful design for determining cause and effect. Random assignment of subjects to research conditions provides strong protection against common threats to internal validity, and inferences about cause and effect can be made with acceptable degrees of confidence. However, quasi-experiments and true experiments that test the Rules are very rare.

The concept of external validity must also be examined in determining the benefits of the present way of studying the Federal Rules of Civil Procedure. According to Cook and Campbell, external validity refers to the degree to which causal research conclusions can be generalized from the particular setting of the study to the setting of interest. Cook and Campbell go on to suggest that one should ask whether the setting, persons, and time of application are similar enough to the research conditions to allow a confident conclusion that the findings of the research can be used in solving the problem at hand. Current rule changes are implemented in the entire federal trial court system and investigated in that context. Thus, the settings of testing and application are the same. Of course, some new persons would be involved in future years and the passage of time between implementation and investigation might bring changed conditions. Nevertheless, it is reasonable to conclude that information about the impact of rule changes developed in this way would have strong external validity. This high external validity is one benefit of current research practices.

The net benefits of current research methods are surely very low, however. The slight internal validity of causal conclusions developed by most current

31. See supra notes 14-16 and accompanying text.
32. See Diamond, supra note 24, at 675-80.

The situation, discussed above in Part III, where some states employ the Federal Rules, others use variations on the rules, and still others have entirely different systems, creates an opportunity to make useful comparisons in studying procedure. Such comparisons would enhance internal validity, but the possible effects of differences in general conditions (e.g., employment, weather, politics) would remain a concern. Cf. Stigler, What Can Regulators Regulate?: The Case of Electricity, in THE CITIZEN AND THE STATE: ESSAYS ON REGULATION 61 (1975) (Claire Friedland collaborating); Etzioni, Does Regulation Reduce Electricity Rates? A Research Note, 19 POL'y SCI. 349 (1986). Also, concerns about external validity, discussed next in the text, would be heightened because of differences between state and federal courts.

34. See Lind, Shapard & Cecil, supra note 33, at 88-92. See also J. MONAHAN & L. WALKER, supra note 27, at 59-62.
35. See Diamond, supra note 24, at 670-71.
36. T. COOK & D. CAMPBELL, supra note 27, at 70-73.
37. Id. at 73-74.
designs provides no safe basis for conclusions about the impact of particular rules.\textsuperscript{38} Without this logically based confidence, the high external validity of the research is virtually meaningless. Thus, the results regarding cause and effect are so uncertain that general applications based on expected impact are unwise.

Estimating the costs of current practices necessarily involves speculation because these very research practices do not permit confident conclusions. This difficulty is particularly severe in the case of error costs because the impact of particular rules and rule changes is unclear under the present research regime. But there are, at least, strong opinions that some Rules may have produced surprising results that have added to the individual and social burden of litigation. For example, the 1966 revision of the class action rule, Rule 23, has been described as “a form of legalized blackmail”\textsuperscript{39} and as used to produce a “Frankenstein Monster.”\textsuperscript{40} The 1983 amendment of Rule 11 has been described as being likely to “further erode public confidence in our system of justice”\textsuperscript{41} and as threatening to “chill vigorous advocacy.”\textsuperscript{42} The comprehensive system of discovery adopted in the original Rules has been said to result sometimes in “a massive and unequalled invasion of privacy . . . .”\textsuperscript{43} Another observer claims that relaxed pleading requirements and broad discovery have produced unintended results: “The theory was that this would prevent pleading from being a ‘game of skill’ and prevent trials from becoming ‘sporting matches’. The practice—in many areas of the law—has been to make discovery the ‘sporting match’ and an endurance contest.”\textsuperscript{44} Almost certainly, some rules approved by policymakers are of poor quality. In any event, there is at present no valid way of knowing which changes are beneficial. In this situation, it is reasonable to assume that while some changes are beneficial, others have the opposite effect.\textsuperscript{45} Certainly, this is the prudent conclusion until there is good reason to conclude otherwise.

The transaction costs of doing the research are also probably high. The cost of printing and distributing a rule change is not great, but the time spent by judges and lawyers learning about the change is probably considerable.\textsuperscript{46}

\textsuperscript{38} See supra notes 29-30 and accompanying text.
\textsuperscript{39} Handler, The Shift From Substantive to Procedural Innovations in Antitrust Suits, 71 COLUM. L. REV. 7, 9 (1971).
\textsuperscript{40} Eisen v. Carlisle & Jacquelin, 391 F.2d 555, 572 (2d Cir. 1968) (Lumbard, C.J., dissenting).
\textsuperscript{41} Weiss, A Practitioner’s Commentary On the Actual Use of Amended Rule 11, 54 FORDHAM L. REV. 23, 27 (1985).
\textsuperscript{43} Kirkham, Complex Civil Litigation—Have Good Intentions Gone Away?, 70 F.R.D. 199, 203 (1976).
\textsuperscript{44} Rifkind, Are We Asking Too Much of Our Courts?, 70 F.R.D. 96, 107 (1976).
\textsuperscript{45} As will be seen, this assumption is very important to the argument that follows. Valid research results might well require changing the assumption, and that change might indicate that a different research program should be adopted. Indeed, it is conceivable that valid research might show that the intuitive judgments of rulemakers are almost always correct. In this event, the research program might be abandoned. At least the process could proceed with confidence and not with gross uncertainty, as it does now.
\textsuperscript{46} For example, if only one out of six American lawyers spent only one hour reviewing a rule change, more than 100,000 hours of time would be invested. See B. CURRAN, K. ROSICH, C. CARSON.
The cost of collecting data under current practices is comparatively low because so little data have actually been collected. The sum of error costs and transaction costs might be described as moderately high. Thus, consideration of both the benefits and costs of current research practices suggests the need to look for alternatives. The amount of valid impact data produced is small and the cost of collecting this information (considering net error cost and transaction cost) is substantial.

B. National Experiments

One important alternative to current practices is a program of unrestricted or national experiments. Such tests would be carried out by randomly assigning cases throughout the nation to use a proposed new rule and then assessing the impact of that rule in those cases as compared with cases litigated under the existing rule. This plan is important to consider because it is the methodological converse of current research practices. In terms of validity, the benefits of such a national experiment would be very high. The use of a true experimental format, which includes random assignment of cases to research conditions, would give the best guarantee that the effects measured were caused by the rule changes tested. Similarly, the external validity of such a national experiment would be very high. The research done in this format would involve the very courts which would implement any change, and while there would be some changes of personnel and passage of time, conclusions made about the effect of proposed changes would likely generalize in some future application to those same U.S. district courts.

The costs of obtaining this increased validity would probably be lower than under current practices. Assuming that some of the innovations are poor in quality, this scheme would quickly identify bad changes and permit confident elimination of those rules, thus sharply reducing the negative effect of repeated use. Of course, this testing regime would delay, for control group cases, the use of a beneficial rule, but the denial would be temporary. Beneficial changes would be quickly (and confidently) identified and given general application. Thus, over time, the total error cost of a true experimental system would probably be much lower than that of current practices. Such national experiments would produce substantial improvement in the Rules of Civil Procedure; accordingly, the court system, in implementing those Rules, would be improved.


47. An alternative method might involve random assignment of federal districts to test or control conditions. This might ease practical problems of conducting the research, although important statistical considerations would be implicated by a sharp reduction in the number of sample units. See T. Cook & D. Campbell, supra note 27, at 354-56. This method would also introduce concerns about forum shopping.


49. See supra notes 36-37 and accompanying text.
The transaction costs of informing all judges and most attorneys about a proposed rule change and the cost of collecting data about every case tried in the U.S. district courts would be high. Thus, the benefits of such a plan would be great but the sum of error costs and transaction costs would be considerable.

C. Restricted Field Experiments

A second important alternative would be to implement similar true experiments in a small number of U.S. district courts. The internal validity of the research for causal information would be high because a powerful research method would be used to collect information. The external validity of the research results would be lower than that of either current practices or a system of national experimentation, because restricted field experiments would enhance the possibility that different settings would affect the application of research results. Any resulting changes would be introduced in many courts not involved in the testing. In sum, the benefits of conducting a restricted field experiment would be considerably greater than the benefits under current practices, though certainly not as great as under a program of national experimentation.

A restricted field experiment would bring important advantages in cost reduction. If the research were conducted in only a small number of U.S. district courts, the cost of subjecting persons to rules of dubious quality would be less than under current practices. But the predicted error cost of a restricted field experiment would be about the same as for a national experiment. Given the assumption that some of the proposed changes would be beneficial, the reduction in sample size would not only sharply reduce the exposure to bad rules, but would correspondingly reduce exposure to good rule changes. Reduction in cost of exposure to a bad rule would presumably be offset by the wide withholding of beneficial changes. Therefore, the chief cost advantage would result from reduced transaction costs. These costs would be much lower than those under current practices or under a national experiment. The cost of promulgating a new rule and of measuring the results of that rule would be sharply diminished in the restricted field experiment format.

In sum, current research practices produce very little benefit with considerable cost. In contrast, a program of national experimentation would produce great benefit with considerable cost. Finally, a program of geographically restricted field experiments would produce substantial benefit at moderate to low cost. Although a program of national experiments is certainly appealing, restricted field experiments are a progressive compromise.

In principle, restricted field experiments should be used to study every proposed change to the Rules, except amendments intended only to clarify existing Rules. The logic of the proposed method does not suggest a criterion for selection of Rules to be tested. However, the legal and ethical
questions discussed below do suggest that it would be desirable to test
changes believed to be generally beneficial—as opposed to changes believed
to favor one group of participants (that is, plaintiffs or defendants). Also, the
practical problems of using the present Rule and a test rule in the same court
(but in different cases) obviously must be considered. Selected projects
should involve clear, comparatively simple differences so that judges and
lawyers can easily follow one rule in a particular case and another rule in
another case.

A recent agenda of the Advisory Committee on the Civil Rules provides
several good candidates for field experiments.\textsuperscript{50} Item B would permit the
"strike method" of jury selection.\textsuperscript{51} This change is believed to be generally
beneficial and could be tested easily by a restricted field experiment.
Permitting the strike method is a clear, simple change. The comment
accompanying this proposal states that "there appears to be no sound
objection to this method," which suggests a particularly uncontroversial
candidate for research.\textsuperscript{52} Item J also is a strong candidate. Item J proposes to
revise Rule 40 to include provisions planned to improve the management of
complex litigation.\textsuperscript{53} Again, this appears to be a generally beneficial change,
and cases could easily be randomly assigned to operate with the present or
proposed Rule 40. Although the proposed changes are substantial, it is
intuitively non-controversial to use different management techniques in
different cases. Agenda item I proposes, essentially, to eliminate the Rule
12(b)(6) motion to dismiss for failure to state a claim.\textsuperscript{54} The accompanying
comment suggests general benefits, and restricted experimental testing would
be simple and uncontroversial.\textsuperscript{55}

V
LEGAL QUESTIONS

Legal objections to the restricted field experiments described above would
likely be based upon a claim that the research plan constituted a denial of
equal protection.\textsuperscript{56} Challenges might come either from litigants randomly
assigned to test conditions, or from litigants in control conditions who were
not permitted to use a test rule. The first careful analysis of the issues

\textsuperscript{50} See Advisory Committee on the Civil Rules, Preliminary Discussion Drafts
(unpublished agenda, meeting of Apr. 7-8, 1988, P. Carrington reporter).
\textsuperscript{51} Id. at 4.
\textsuperscript{52} Id. at 5.
\textsuperscript{53} Id. at 57-60.
\textsuperscript{54} Id. at 53.
\textsuperscript{55} Strategies for gaining public acceptance of the use of true experiments in policy research are
\textsuperscript{56} The due process language of the fifth amendment would provide specific authority for such
a challenge to federal court action, but the entire body of federal equal protection jurisprudence
would apply. See Bolling v. Sharpe, 347 U.S. 497, 499-500 (1954); Steward Machine Co. v. Davis, 301
("[T]he due process clauses of the fifth and fourteenth amendments also have been held to yield
norms of equal treatment indistinguishable from these of the equal protection clause.").
LAW
AND
CONTEMPORARY
PROBLEMS

discussed in this section was done in 1959 by Professor Hans Zeisel and his co-authors Professor Harry Kalven and Bernard Buchholz. In their book, *Delay in the Court*, in a chapter titled “The Case for the Official Experiment,” the three authors briefly described the advantages of the true experiment, and then turned to examination of available precedents relating to experiments with court rules. Although they report no case law dealing with an instance of random assignment, their wide-ranging survey concludes, “[i]n the end then it would appear that the temporary differentiating of the official experiment would not necessarily be discriminatory, and would serve the high purpose of making more secure the knowledge on which the administration of justice is based.” While this conclusion appears to have been sound at the time they wrote, the question arises whether developments in the law since 1959 have changed the situation.

Proposed experimentation with the Federal Rules of Civil Procedure would be examined properly by a court under the “rational basis” test for determining whether a statute violates the equal protection guarantees of the Constitution. The “strict scrutiny” test would not apply because neither a suspect classification nor a fundamental right would be involved. Also, an “intermediate” test would not apply to the proposed research plan because the classification would not involve gender or illegitimacy. The enterprise of empirical research is rational, provided that the method of the investigation establishes a logical basis to use the results to improve the court system, an important public service. As discussed above, a restricted field experiment would furnish a strong logical case for adopting proposed Federal Rules of Civil Procedure that prove to be beneficial under these conditions.

---


59. The courts have allowed facially neutral classification of groups in legislative acts so long as the classification or burden placed upon the group is rationally related to a legitimate governmental objective. The test is highly deferential to the legislative body—the presumption being that the legislature acted within its power and that the equal protection guarantees were violated “only if the classification rests on grounds wholly irrelevant to the achievement of the state’s objective.” McGowan v. Maryland, 366 U.S. 420, 425 (1961). See Wilkinson, *The Supreme Court, The Equal Protection Clause, and the Three Faces of Constitutional Equality*, 61 VA. L. REV. 945, 950-54 (1975); L. TRIBE, *supra* note 56, at 1439-43.

60. If the challenged classification is “suspect” (e.g., based on racial characteristics) or directly impinges a fundamental right, the Court permits the classification only if it serves a compelling legislative interest. Suspect classifications have been largely confined to matters of race, alienage, and national origin. Fundamental rights have been limited to those established by the Constitution (e.g., voting, travel, or association). See Wilkinson, *supra* note 59, at 951; L. TRIBE, *supra* note 56, at 1451-54.

61. Classifications triggering closer scrutiny than the “rational basis” test are not uniform either in the formulation of the standard for review or in application. A useful elaboration of the test in this amorphous middle category is that such classifications “must serve important governmental objectives and must be substantially related to achievement of those objectives.” Craig v. Boren, 429 U.S. 190, 197 (1976).
The random assignment of cases for research purposes has received some attention since 1959. These research conditions have already been employed at least once by an American trial court to test civil rule changes without any reported objection. According to Connolly and Planet, Kentucky's Campbell County Superior Court conducted a two-year-long research project in which most classes of civil cases were “randomly and evenly distributed between the two dockets.” Different civil rules were used for each of the two dockets. The difference in the two sets of rules was substantial: Motion practice was changed, discovery was limited, and pretrial was strictly monitored. Connolly and Planet do not report any legal objection to the use of random assignment and there are no reported Kentucky decisions relating such an objection. Thus, to this limited degree, a precedent exists for the current proposal.

In addition to this Kentucky project—particularly relevant because trial court civil rules were tested—at least five other limited precedents exist. During 1978, the United States District Court for the District of Connecticut permitted the use of virtually true experimental method to test the effect of mandatory, nonbinding arbitration in certain types of civil actions. No reported opinion reveals an objection to this procedure. Also, the United States Court of Appeals for the Second Circuit has twice permitted use of true experimental method to test the effect of several appellate case management practices. There is no reported record of objection. In 1987, the United States Court of Appeals for the District of Columbia Circuit ordered random assignment of most classes of civil appeals to mandatory mediation. No objections have been reported. Finally, one of Pennsylvania's two

---

62. Connolly & Planet, Controlling the Caseflow—Kentucky Style, 21 Judges’ J., Fall 1982, at 8. The Campbell County (state) project also is described briefly in Note, Economical Litigation: Kentucky's Answer to High Costs and Delay in Civil Litigation, 71 Ky. L.J. 647, 649-52 (1983). Although not involving random assignment, the idea of “trying out” changes in civil procedure on a limited basis also has been used in California. See Epstein, Reducing Litigation Costs for Small Cases, 20 Judges’ J., Spring 1981, at 8. The program was mandated by statute, Cal. Civ. Proc. §§ 1823-1833.2 (West 1983). Section 1823 included the following finding:

The Legislature further finds and declares that there is a compelling state interest in the development of pleading, pretrial and trial procedures which will reduce the expense of litigation to the litigants and there is likewise a compelling state interest in experimentation on a small scale with new procedures to accomplish that result before those procedures are adopted statewide. Therefore, the provisions of this part are added to this code to provide a means of experimentation with procedural innovations to reduce the cost of civil litigation.

63. Connolly & Planet, supra note 62, at 11.

64. Id. at 13.

65. See E. Lind & J. Shapard, supra note 6, at 18-19.


intermediate appellate courts randomly assigned civil cases to mandatory settlement conferences, also without reported objections.\textsuperscript{68}

The leading case discussing the random assignment of subjects for research purposes is \textit{Aguayo v. Richardson}.\textsuperscript{69} In that case, plaintiffs challenged, on equal protection grounds, a plan to carry out a research project about welfare benefits. According to Judge Friendly's opinion, the program involved the random assignment of various political subdivisions of New York state to different welfare regimens. The Second Circuit denied the plaintiffs' challenge, holding that "the Equal Protection clause should not be held to prevent a state from conducting an experiment designed for the good of all, including the participants . . . ."\textsuperscript{70} Although not involving a test of civil procedure rules, \textit{Aguayo} is similar in importance to the proposed experiments because both \textit{Aguayo} and the proposal involve government-operated research projects which would entail random assignment to test conditions. One difference between that case and the proposal is that the \textit{Aguayo} case involved random assignment of whole political subdivisions, while the proposal involves random assignment of individual cases to test or control conditions. This distinction is not important, however, because the different treatment is, in both instances, a function of chance employed to rationalize an investigation method.\textsuperscript{71}

A second specific objection might focus on disparity of treatment by location. Geographically based disparity is inherent in the proposal of restricted studies. In \textit{Kimbrough v. Holiday Inn}\textsuperscript{72} defendants challenged, on equal protection grounds, a program instituted to test the use of compulsory

\begin{itemize}
\item \textsuperscript{68} Steelman & Goldman, \textit{Preargument Settlement Conferences in State Appellate Courts}, 10 STATE CT. J., Fall 1986, at 4.
\item \textsuperscript{69} 473 F.2d 1090 (2d Cir. 1973), cert. denied, 414 U.S. 1146 (1974). Although apparently not involving random assignment, important support for evaluation research was stated in California Welfare Rights Org. v. Richardson, 348 F. Supp. 491, 497 (N.D. Cal. 1972). ("It goes without saying that experiments upon the most basic living conditions of human beings who are to a very large extent totally dependent upon governmental programs for their existence, is a most serious matter. The only thing more serious than an experimental program that does not produce beneficial results, is the non-experimental, standard program statutorily mandated which so fails.").
\item \textsuperscript{70} 473 F.2d at 1109.
\item \textsuperscript{71} Although not stated in a judicial opinion, it is nevertheless noteworthy that Judge Posner clearly appreciates the importance of random assignment. Posner wrote:
\begin{quote}
My second criterion for a proposed alternative to the conventional trial was that its effects be verifiable. This ideally would require that cases be assigned randomly for summary jury trial, with other cases receiving the usual treatment. And by random, I mean random; I mean by use of a table of random numbers. Then the summary jury trial cases could be compared with a proper control group and, provided there were enough cases in each group to allow statistical comparison, we would know whether there were any differences and, if so, what they were.
\end{quote}
Posner, \textit{supra} note 4, at 374.
\end{itemize}
arbitration in civil suits. The defendants in that case claimed that the arbitration program treated them differently than litigants in similar districts that were not part of the program. The court acknowledged that "litigants in three federal districts arbitrate prior to trial while other litigants nationally do not," but held that there was a rational basis for the arbitration research. Similarly, in McGlothen v. Department of Motor Vehicles a California court held that a demonstration and research project limited to four counties did not result in the denial of equal protection of the laws with respect to cases disposed of in other California counties. The rule tested in McGlothen provided that the license of a motorist convicted a second time for driving under the influence of alcohol would not be revoked, provided the motorist participated successfully in a rehabilitation program. The four plaintiffs were all motorists who had suffered second convictions in counties not part of the demonstration project. The trial court enjoined the Department of Motor Vehicles from revoking their drivers' licenses, provided they too participated in a rehabilitation program. The California appellate court reversed the action of the trial court and held that the geographical disparity did not violate the motorists' right to equal protection.

Thus, it appears that the 1959 view of Zeisel, Kalven, and Buchholz remains sound. Indeed, legal developments since that time have increased support for responsible governmental use of true experimental method. Since 1959, at least one trial court has used random assignment to test civil rule changes and at least four other courts have used true experimental method to test important aspects of civil case management. Furthermore, there is now an appellate precedent approving random assignment in a major governmental research project.

VI
Ethical Questions

In 1978, Chief Justice Burger appointed a committee to consider ethical questions that might be presented by proposals to do empirical research in the federal courts. The committee issued a report in 1981, which suggests a
series of tests appropriate for analysis of specific research projects. Four basic questions are proposed, three dealing with ethical issues and a fourth concerning the authority of officials to undertake research. This fourth question is discussed in the section immediately following.

The first test proposed by the committee asks, essentially, whether there is sufficient practical justification for the proposed research. As discussed above, there is a need to evaluate and improve the civil procedure system. There is uncertainty about the effect of most proposals for change and restricted field experiments offer a reliable means to supply this information. Research results would indicate a course of action for procedural reform, a goal clearly within the public interest.

The committee's second test asks what research design would meet the demonstrated need. Again, the discussion above shows that restricted field experimentation is responsive to the need. Although unrestricted field experimentation would probably produce the most useful results, a restricted program would produce acceptable results with modest cost and inconvenience.

The third test examines ethical problems that a particular proposal might present, incorporating two principles, "equal treatment" and "respect for persons," as deserving particular consideration. Here, the proposal of random assignment of cases to different procedural treatments appears to present the most serious questions, given the principle of equal treatment. In an extended discussion of this issue, the committee suggests consideration of six discrete factors: (1) significance of the interests affected; (2) extent of the difference between treatments; (3) disparity with standard treatment; (4) differences in qualifications of subjects; (5) potential for harm and benefits; and (6) voluntariness of participation.

Review of these factors does not suggest any major barrier to the proposal. Indeed, research about civil procedure may be particularly appropriate according to these criteria. The interests affected are important but probably do not determine outcome. In any event, loss of personal liberty is virtually never at stake. The probable differences in treatment are comparatively slight, and the differences are substituted for identical treatment, a favorable situation, according to the committee. Furthermore, any test condition would

78. Experimentation in the Law: Report of the Federal Judicial Center Advisory Committee on Experimentation in the Law (1981). Reference to the report invites a cautionary comment. Unfortunately, the committee chose to use the term "experiment" to refer to "a test of a new concept or program" rather than to a "randomized experimental design." Id. at 2 n.2. This puzzling choice can confuse the unwary reader. Some participants in the project rejected this view, at least for their own contribution. See Lind, Shapard & Cecil, supra note 33, at 9 n.37. The committee's views have been controversial. According to Judge Posner, "The characteristic hostility of law-trained persons for experimental studies of law is illustrated by" the committee's report. See Posner, supra note 4, at 368 n.4.
79. Experimentation in the Law, supra note 78, at 7.
80. Id. at 11.
81. Id. at 15.
82. Id. at 25-27.
83. Id. at 31-40.
be intended to benefit subject cases, also recognized as a favorable factor. The circumstance that cases would be assigned to different treatments without regard to their particular characteristics and on a mandatory basis is viewed negatively by the committee—although the use of true experimental design requires these conditions. On balance, analysis of the six factors reveals no categorical bar to the proposal.

VII
IMPLEMENTATION

The committee’s fourth test focused on the need for authority to undertake research and on the related need for procedures to guide the course of such research. Thus, a question arises: What is the best way to implement a program of restricted field experiments for proposed federal rule changes? Perhaps the most promising course would be to revive an apparently abandoned proposal to amend Federal Rule 83 to expressly permit evaluation of procedural alternatives by local rule.

In 1983, the Civil Rules Advisory Committee, as part of a recommended general revision of Rule 83, proposed to add the following sentence to the Rule: “When authorized by the judicial council, a district Court may adopt on an experimental basis for no longer than two years a local rule that may not be challenged for inconsistency with these rules, after giving appropriate public notice and an opportunity to comment.” The Advisory Committee Note provided, in part, as follows:

The second paragraph of amended Rule 83 adds a new provision allowing experimental local rulemaking, including provisions that may be inconsistent with the Federal Rules. This is designed to be a useful mechanism for carefully testing and evaluating procedural proposals to see if they are better tailored to local conditions than existing practices or alternative proposals. Well-conceived experiments also can provide an experience base that might prove useful for national rulemaking. See Flanders supra at 213. The new language requires prior authorization by the judicial council because of the sensitivity of experimenting in a way that may be inconsistent with the Federal Rules and is subject to the same public notice and comment procedure required by the first sentence of the rule for nonexperimental local rulemaking. The experimentation also is limited to a period of two years. If an experimental rule is to be made permanent, the full procedure for amending the Federal Rules must be followed.

84. Id. at 35-36.
85. Id. at 38-40.
86. Id. at 67-76.
87. Early support for this type of change was given in Flanders, Local Rules in Federal District Courts: Usurpation, Legislation, or Information?, 14 Loy. L.A.L. Rev. 213, 276 (1981). He recognized that changes in local rules might produce information useful to shaping general amendments. Id. at 269. The proposed amendment to the Rule is discussed in Roberts, The Myth of Uniformity in Federal Civil Procedure: Federal Civil Rule 83 and District Court Local Rulemaking Powers, 8 U. Puget Sound L. Rev. 537 (1985).
89. Id. at 370.
90. Id. at 373. The reference to Flanders in the quotation is fully cited supra note 87 (footnote omitted).
This proposed change in Rule 83 was not ultimately adopted. The former proposal provides a useful starting point for implementation of restricted field experiments, although some changes should be made. The proposal does correctly address the obvious problem with using current Rule 83: Test proposals would be inconsistent with the Rules, and current Rule 83 prohibits such inconsistency. Also, the former proposal correctly provides for ample judicial supervision of projects, obviously a key to the implementation and ultimate success of the experiments.

The shortcoming of the proposal (and the related Committee Note) is the omission of clear reference to random assignment of cases to test conditions. Although the word "experimental" is used in the proposed Rule change, with similar language in the Committee Note, the authorization to adopt "a local rule" might suggest a requirement of general application within the district. Therefore, at a minimum, a future Committee Note, in explaining the change, should include express approval of the random assignment of cases to test conditions. Indeed, authorization is surely not worth obtaining if random assignment is not clearly permitted.

Another shortcoming is the two-year limit. In general, the proposed strict limitations on experimentation seem appropriate: judicial council approval; public notice and comment; time limitation. However, the two-year limit is not realistic. Although many projects could be completed in that time, empirical research is often very time-consuming, and many good projects would require more than two years. Either a flexible time provision or a longer period should be established.

VIII
Conclusion

Now is the time to make the process of federal civil rulemaking fit for future challenges. Rulemakers should adopt an official plan to collect valid information about the probable impact of proposed rule changes before those changes are put into general use. The technique of restricted field experiments can be regularly employed to perfect federal civil rules. The public interest would be well served by such a step; moreover, there are no legal or ethical barriers to taking this important action. A good technique for implementing such a program is at hand, and institutions able to carry out the

---

92. A committee note might also expressly recognize a role for the Civil Rules Advisory Committee in planning research. National coordination would be essential and local initiatives should be discouraged because of the risk that failure would jeopardize the entire program. Also, participation by the Federal Judicial Center should be facilitated. The Center, now a mature and respected institution, is available to assist in planning and carrying out the proposed research. The Center's function is explained in 28 U.S.C. § 620 (1968), which provides, in part, that the Center shall "conduct research and study of the operation of the courts of the United States . . . ." The Rand Corporation's Institute for Civil Justice, a respected private organization, might be included on a consulting basis or employed to carry out specific projects.
work exist. The maturity of the Rules and the development of social science permit this advance in perfecting civil procedure.