Randomized Control Trials in the United States Legal Profession

D. James Greiner and Andrea Matthews

Harvard Law School, Cambridge, Massachusetts 02138; email: jgreiner@law.harvard.edu, amatthews@jd15.law.harvard.edu

Annu. Rev. Law Soc. Sci. 2016. 12:295–312

First published online as a Review in Advance on April 29, 2016

The *Annual Review of Law and Social Science* is online at lawsocsci.annualreviews.org

This article's doi: 10.1146/annurev-lawsocsci-110413-030732

Keywords

randomized control trial, field experiment, legal epistemology, legal profession

Abstract

We assemble studies within a set that we label randomized control trials (RCTs) in the US legal profession. These studies are field experiments conducted for the purpose of obtaining knowledge in which randomization replaces a decision that would otherwise have been made by a member of the US legal profession. We use our assembly of approximately 50 studies to begin addressing the question of why the US legal profession, in contrast to the US medical profession, has resisted the use of the RCT as a knowledge-generating device.

INTRODUCTION

In 1933, two college professors and a legal academic published what to our knowledge is the first study of field data derived from a randomized[1] data-generating process in a legal setting (Gaudet et al. 1933). The date is significant. Only two years earlier, three physicians (Amberson et al. 1931) published what at least one scholar describes as the first study making formal use of randomization in medicine (Meldrum 2000). One year earlier, the US Supreme Court referred to a lay person's need for assistance in understanding "the science of law" [*Powell v. Alabama* (1932)]. Eighty years ago, then, one might have been hard pressed to guess which of two quintessential United States professions, law or medicine, would undergo a transformational epistemological change such that the randomized control trial (RCT) would become the gold standard for the evaluation of interventions provided to patients or clients.

We all know what happened. Although the transformation in medicine has (rightly) been criticized as failing to extend far enough (Finkelstein & Taubman 2015, Parker 2005), there is no comparing the amount or quality of the objective, scientific evidence available to practitioners in the two professions. And despite a compelling literature on failures within the medical profession to incorporate RCT findings into day-to-day practice (e.g., Boissel 1989), there is no comparing the evidentiary basis for the standard of care as between medical and legal professionals. Practitioners of medicine chose to transform their profession into a science. Practitioners of law did not.

We think this situation is bad, on the law side. We think that the United States would be a better place if the legal profession were less hostile to objective, rigorous, scientific evidence about causation and the effectiveness of interventions. We think all this particularly true in the areas of (*a*) interventions for individuals unable to hire attorneys to address their legal problems, civil or criminal, transactional or litigation, and (*b*) the construction and administration of adjudicatory systems. These two arenas, unlike those in which legal professionals and judges compete for business, lack the discipline that markets can sometimes impose on inefficient or wasteful practices. We acknowledge the critical role that theoretical research (Albiston & Sandefur 2013), doctrinal analysis, and observational studies play in the construction of a foundation for addressing socioeconomic and legal problems, and we believe that there are many methodologies that must be deployed if a field is to generate that foundation. Nevertheless, under the present state of our understanding of methodology, no field can claim to be evidence based without a central role for the RCT as a means of accumulating knowledge about what works and what does not.

We defend none of these claims in this article (see Green & Thorley 2014 for the beginnings of such a defense). Nor do we follow here the path of a traditional review paper in reviewing the literature of a domain, outlining the domain's contours and debates, and speculating about the domain's future. We make these choices because, at present, none of these things exist with regard to RCTs in the US legal profession. There are no recognized papers in this domain, no canonical studies, no contours of debate or contesting schools of thought, no internally defined best practices, and few publications proposing agendas for the future. At present, there is no domain to review.

Thus, in this article, we attempt to begin a conversation on the following subject: Why is there no domain to review? What can be done to overcome the US legal profession's current disdain for the RCT? Relatedly, why did medicine transform itself into a science, while law did not?

[1] Admittedly, the study was not formally randomized in the sense of assigning study units to treatments according to coin flips, dice rolls, a random number table, a computer's random number generator, or some other recognized method of producing an unconfounded treatment assignment process. Instead, judges were assigned to defendant sentencing proceedings according to a rotating wheel, i.e., a list of judge names that repeated (Gaudet et al. 1933, p. 813). As noted below, absent some reason for doubt, we credit a rotating wheel as producing an unconfounded treatment assignment.

This article is our first step toward beginning to answer these questions. Here, we catalog and briefly analyze those RCTs that do exist in the US legal profession. As we explain further, our analysis of these studies focuses not on their scientific content but rather on what they can teach us about the questions outlined above. The intensity of the US legal profession's resistance to the RCT is such that, viewed individually, each law RCT appears to be a unicorn, a magical creation with no origin story that appears briefly in a larger setting and then fades away. We expect that analysis of all RCTs in the US legal profession will reveal generalizable lessons about this pattern of spontaneous emergence and later obscurity.

We organize this paper as follows: We first define what we mean by RCT in the US legal profession. We then provide the results of our three-year search for such studies. Second, we defend our claim that the US legal profession is resistant to the RCT. On this point, our primary evidence is the paucity of such studies in law as compared with medicine or even with less quintessentially professional endeavors. To illustrate further, we provide three anecdotes of hostility researchers encountered when they were able to mount randomized field operations in the US legal profession. Finally, we formulate questions for future investigation into the source of the US legal profession's resistance to the RCT, and we hypothesize possible answers to those questions, focusing in particular on answers that might suggest ways in which the present situation could be changed.

OUR INCLUSION CRITERIA AND A REVIEW

Here we state our inclusion criteria for the set of RCTs in the legal profession and discuss the results of our three-year search for such studies. To clarify, we do not suggest that studies outside our set of interest are less scientifically informative or less worthy of admiration. We focus on the set of studies defined below because we are interested in the questions articulated above, and because we suspect that the key to answering those questions lies in the US legal profession's prevailing epistemological norms and professional identity. We have defined our set of studies so as to encompass those research efforts that most directly challenge how the legal profession thinks of itself.

RCTs in the US Legal Profession

We have five inclusion criteria for our review. First, we are interested in studies that took place in the United States. We impose this criterion because, in contrast to many other nations, the United States has chosen to imbue members of its legal profession with unusual attributes, including self-regulation; practitioner autonomy; state-assisted exclusion of competitors; immunity from certain legal doctrines (such as presumptive dormant Commerce Clause invalidation of restrictions on cross-border competition); and, at least in the case of elite members of the profession, wealth, status, and power (Gorman & Sandefur 2011). As compared with other fields of endeavor, lawyers dominate the US governmental structure. Much may be changing, in the United States and elsewhere (Susskind 2010), but until that change is further along, the US legal profession represents an unusually difficult target for transformation. If we can render the legal profession less hostile to rigorous evidence on causation in the United States, we may have luck doing so elsewhere.

Second, we are interested in field studies of interventions, in research efforts that involve interventions applied to actual cases, or clients, or attorneys, or other real-world objects as randomization units. Third, we are interested in studies involving a randomized assignment mechanism or a scheme that, given the setting in which it operates, can be expected to produce unconfoundedness. That means that each object of the study (the unit) must be assigned to a treatment condition (e.g., treated or control) in a manner unrelated to its outcome values (see Imbens & Rubin 2015 for definitions and explanations of unfamiliar terms and concepts). Assignment mechanisms involving

coin flips, dice, random number tables, and computers certainly qualify. In some legal settings, treatment assignment is accomplished by means of some rotational device, such as odd/even case numbers (Goldman 1978). As suggested above, absent some reason to fear manipulation of or departure from these mechanisms or a periodic trend in the data (e.g., Chilton & Levy 2015, Thorley 2015), we credit them for producing unconfoundedness. Note also that these inclusion criteria exclude studies that randomly select units for observation, such as sample surveys (and the random selection of class action cases for bellwether trials).

We impose the field and randomized intervention criteria because, scientifically, field experiments enjoy heightened internal validity and persuasive power (Green & Thorley 2014). Sociologically, field experiments require that members of the US legal profession engage with scientific endeavor in a way that lab (e.g., Elek et al. 2012) or online (e.g., Robinson & Yokum 2012) studies do not.

Fourth, we are interested in studies that are in the legal profession in the sense of involving the replacement by randomization of a decision-making process otherwise undertaken by lawyers and judges. An example may clarify. In our search, we found studies in which judges followed a randomized assignment mechanism in sentencing (or not sentencing) criminal defendants to some program such as counseling or intensive monitoring (e.g., Labriola et al. 2008). We found other studies in which a judge sentenced defendants to conform to the dictates of some office whose members are not in the legal profession, e.g., to probation, and that office followed a randomized assignment mechanism in deciding which defendants would be required to undergo some program, such as counseling or intensive monitoring (e.g., Petersilia & Turner 1991). Assuming all other conditions are met, the first studies are in our set of interest, and the second are not. Judges are frequently lawyers and are considered members of the legal profession. Probation officers frequently are not lawyers, and their work is not ordinarily considered practicing law. Thus, our focus is not on whether some decision is legal in some sense. Our review of the literature suggests that nonmembers of the legal profession in some jurisdictions make decisions that, in other jurisdictions, are made by members of the legal profession. Rather, our focus is on who, but for the randomization, would have made the decision.

We impose the requirement of "in the legal profession" because our own experiences and our reading of the literature suggest that it is the lawyers and judges, as opposed to the nonlawyers operating within the field of US law, who are resistant to the RCT in particular. The US legal profession is structured on the premise that only a lawyer knows how to address legal problems, and that the only person a lawyer should listen to for ideas on how to address legal problems is another lawyer [e.g., Mass. Gen. Law ch. 221 §46A (the Massachusetts unauthorized practice of law statute) and ABA Model Rule 5.4 (prohibiting US lawyers from using certain business structures so as to preserve the "professional independence of the lawyer")]. The profession reifies professional judgment and the personal experiences of its practitioners. Thus, a primary challenge to mainstreaming the RCT in the law is persuading legal professionals, meaning lawyers and judges, to replace their judgments with a randomized assignment mechanism for the duration of a study. Another is persuading legal professionals to view RCT results as useful evidence when, e.g., making judgments on how to administer adjudicatory systems or structure the delivery of services. So we focus on the decisions of legal professionals, an approach distinct from that of Green & Thorley (2014), which for the convenience of a literature search seeks field studies using randomization that appear in prominent publications, and from that of Abramowicz et al. (2011), which focuses on legal rules.

Fifth, we are interested in studies in which the reason for randomizing was to generate knowledge. There are a variety of settings in which legal actors use randomization (or a rotational system, e.g., Anderson & Heaton 2012) for reasons other than to learn something, such as recruiting good

young attorneys (Abrams & Yoon 2007), allocating a scarce resource (Cohen 2013), maintaining the appearance of neutrality, or conforming to the politics of the judicial selection process (Samaha 2009). Such studies can produce evidence of the causal effect of the thing randomized (e.g., the judge, the attorney, the type of attorney) and, depending on one's tolerance for modeling and other assumptions, may produce evidence of the causal effect of variables that covary with the thing randomized. None of these studies fall within our parameters. Instead, we are interested in studies in which legal professionals temporarily gave up their decision-making power because they acknowledged at least some doubt as to the effectiveness of some intervention.

The reasons for our "acquisition of knowledge" and "in the legal profession" criteria are similar. To reiterate, we do not assert that a study that takes advantage of randomization implemented for a purpose other than knowledge acquisition is in some way suspect or less worthy than one in which a scholar has persuaded members of the legal profession to engage in a research effort. Our point is that there will never be enough of the former kind of study to provide enough opportunities to gain the knowledge needed to transform law into an evidence-based field.

We close this subsection by clarifying items that are not a part of inclusion criteria. We do not require that the field operation be properly designed or conducted or that the data be properly analyzed. We found numerous statistical errors in the studies in our set, some of them fundamental. But our focus is less on the knowledge those operations produce than on the need to persuade, and process of persuading, members of the legal profession to engage in RCTs. We also do not require blinding. Although we agree that ordinarily, the more layers of blinding, the better, we have found it difficult to blind legal interventions.

Our Search

In this subsection, we describe our three-year effort to find RCTs in the US legal profession that meet our inclusion criteria. Our search unfolded in two phases, the first designed to be an orderly sweep of the widest array of publications available to us, and the second a more ad hoc search based on references provided by existing articles, by colleagues, or by other one-off referrals. The initial sweep surveyed three literature aggregators and one preexisting compilation of studies known to one author in the field of criminology. The aggregators were EBSCO, a major purveyor of social science research; JSTOR, an archive of leading academic journals in a wide range of disciplines; and finally the Harvard University Library's Google Scholar, which searches every full-text database accessible by the University. Within these aggregators, we searched for variants of the word "randomized" by truncating the term to "random" and adding a proximity search to the terms law, legal, regulation, and ordinance. Where the results numbered in the hundreds, each article was scanned quickly and manually for conformation with our criteria. Where the results numbered in the thousands, the first 200 were scanned. The criminology compilation was generously shared by the Urban Institute's Justice Policy Center, one author's former employer. That compilation—the subject of hundreds if not thousands of hours of dedicated searching for field studies in criminology—was then reviewed specifically for RCTs, which were then narrowed by our inclusion criteria.

The initial sweep yielded some, but ultimately less than half, of the studies included in this review. We found most of the studies in our set of interest through haphazard means: word of mouth, or citations in other RCTs, or references in literature reviews unrelated to the current effort. We take this fact as further evidence of our assertion in the beginning of this article that, at present, there is no developed domain of RCTs in the US legal profession. In addition, the way in which we found most of our references, and the scattered nature of their publication (discussed below), makes us wary of asserting that we found all studies in our set of interest.

A Review

In this subsection, we review and analyze the results of our three-year search for RCTs in the US legal profession. The relevant references are set out in a footnote[2] and appear in **Table 1**. We find this set of studies revealing in several respects. First, RCTs in the US legal profession are more numerous than we anticipated. An exact count requires resolution of an unproductive definitional debate because some field operations generated more than one publication, whereas others involved more than one treatment contrast. But however we counted, we found approximately 50 RCTs that met our inclusion criteria. That number is pathetic when compared with the number of RCTs produced in medicine, or even in social science areas related to law, such as criminology (Farrington 1983, Farrington & Welsh 2005). But it is far more than enough to rebut the argument that the randomized field experiment is impossible in law (Engel 2013). Thus, the US legal profession's resistance to the RCT does not stem from a lack of useful RCT examples. In a profession whose thinking is dominated by precedent, precedents exist.

Second, the studies in our set were difficult to locate. They appeared in economics journals (e.g., Abrams & Rohlfs 2011), National Institute of Justice reports (e.g., Baker & Sadd 1981), state court administrative office technical reports (e.g., Clark et al. 1995), think/research tank research reports (e.g., Davis et al. 2000), specialized legal journals (e.g., Deschenes et al. 1995), psychiatry journals (e.g., Ditman et al. 1967), Federal Judicial Center reports (e.g., Eaglin 1990), "law and" journals (Seron et al. 2001), criminology journals (e.g., Goldkamp & White 2006), monographs (e.g., Goldman 1980), public policy evaluation journals (e.g., Gottfredson et al. 2005), generalist state government publications (e.g., Hanson & Hersey 1991), individual professors' websites (e.g., McEwen 1992), psychology journals (e.g., Rudd et al. 2015), books (e.g., Stapleton & Teitelbaum 1972), and National Center for State Courts reports (e.g., Steelman et al. 1986). Only a handful of the studies in our set were published in traditional law reviews (Ayres et al. 1963, Greiner 2011). Thus, it is possible that the extent of the RCT work in the US legal profession has until now been underappreciated.

Third, the subject areas encompassed by these studies are broad, extending from alternative dispute resolution (ADR) (e.g., Goldman 1978), to the information judges receive and consider at bail (e.g., Ayres et al. 1963), to different models of legal services provision (Seron et al. 2001), to rules surrounding jury deliberations and instructions (Heuer & Penrod 1994a), to the effectiveness of different modes of adjudication (Gottfredson et al. 2003), and to various other areas. Studies in our set effectively randomized incarceration (Ayres et al. 1963, Schneider 1986); indeed, studies outside our set directly randomized incarceration (Berecochea et al. 1973) and effectively randomized the likelihood of a death sentence (Anderson & Heaton 2012). In only one sense were the studies we found narrow in focus: We found no study meeting our inclusion criteria outside of the litigation context. Thus, although we do not attempt to delineate a set of legal interventions that are categorically eligible or ineligible for randomized study (we doubt such a delineation is possible), by reference to studies already done, a vast array of legal subject areas appear to be good candidates for RCTs.

[2] Abrams & Rohlfs 2011; Ayres et al. 1963; Baker & Sadd 1981; Clark et al. 1995; Colbert et al. 2002; Davis et al. 2000; Deschenes et al. 1995; Ditman et al. 1967; Eaglin 1990; Feder & Dugan 2002; Goldkamp & Goddfredson 1984; Goldkamp & White 2006; Goldman 1978, 1980; Gottfredson & Exum 2002; Gottfredson et al. 2003, 2005, 2006; Greenwood & Turner 1993; Greiner & Pattanayak 2012; Greiner et al. 2012; Hannaford et al. 2000; Hanson & Hersey 1991; Harrell et al. 2000; Heuer & Penrod 1988, 1989, 1994a,b; Kakalik et al. 1996; Kobbervig 1991; Labriola et al. 2005, 2008; Mandell & Marshall 2002; Margolin 1970; Marlowe et al. 2003; McEwen 1992; McGarrell et al. 2000; Patridge & Lind 1983; Pearson & Thoennes 1984; Rosenberg 1964; Rudd et al. 2015; Schneider 1986; Seron et al. 2001; Stapleton & Teitelbaum 1972; Steelman et al. 1986; Thoennes 2001; Tomkins et al. 2012; Ursa Inst. 1984; Weitzman 1983; Zuberbuhler 2001.

Table 1 Chronological list of randomized control trials in the US legal profession

Reference	Intervention	Civ/Crim	Pub. Type
Ayres et al. (1963)	Bail information	Crim	Law Rev.
Rosenberg (1964)	Pretrial conference	Civ	Mono.
Ditman et al. (1967)	Probation	Crim	Psychiatry J.
Margolin (1970)	Counseling	Civ	Spec. Law J.
Stapleton & Teitelbaum (1972)	Legal representation	Crim	Mono.
Goldman (1978, 1980)	Settlement conference	Civ	Law Rev. & Mono.
Baker & Sadd (1981)	Diversion program	Crim	NIJ Rep.
Patridge & Lind (1983)	Settlement conference	Civ	FJC Rep.
Weitzman (1983)	Legal representation	Crim	NIJ Rep.
Goldkamp & Goddfredson (1984)	Bail guidelines	Crim	NIJ Rep.
Pearson & Thoennes (1984)	Mediation	Civ	Spec. Law J.
Ursa Inst. (1984)	Legal representation	Crim	NIJ Rep.
Schneider (1986)	Restitution	Crim	Crim. J.
Steelman et al. (1986)	Settlement conference	Civ	NCSC Rep.
Heuer & Penrod (1988)	Jury regulation	Both	Spec. Law J.
Eaglin (1990)	Preargument conference	Civ	FJC Rep.
Hanson & Hersey (1991)	Mediation	Civ	State Gov. Pub.
Kobbervig (1991)	Mediation	Civ	State Gov. Pub.
McEwen (1992)	Alternative dispute resolution	Civ	Web
Greenwood & Turner (1993)	Delinquency program	Crim	Crim. J.
Heuer & Penrod (1994b)	Jury regulation	Both	Spec. Law J.
Heuer & Penrod (1994a)	Jury regulation	Both	Spec. Law J.
Clark et al. (1995)	Mediation	Civ	Tech. Rep.
Deschenes et al. (1995)	Drug court	Crim	Spec. J.
Kakalik et al. (1996)	Mediation	Civ	Think Tank
Davis et al. (2000)	Batterer treatment	Crim	Think Tank
Hannaford et al. (2000)	Jury regulation	Civ	Spec. Law J.
Harrell et al. (2000)	Drug program	Crim	Tech. Rep.
McGarrell et al. (2000)	Restorative justice	Crim	Tech. Rep.
Seron et al. (2001)	Legal representation	Civ	Spec. Law J.
Zuberbuhler (2001)	Mediation	Civ	Spec. Law J.
Thoennes (2001)	Mediation	Civ	Tech. Rep.
Colbert et al. (2002)	Legal representation	Crim	Law Rev.
Feder & Dugan (2002)	Counseling	Crim	Spec. Law J.
Gottfredson & Exum (2002)	Drug court	Crim	Crim. J.
Mandell & Marshall (2002)	Mediation	Civ	Tech. Rep.
Gottfredson et al. (2003)	Drug court	Crim	Crim. J.
Marlowe et al. (2003)	Status hearings	Crim	Crim. J.
Gottfredson et al. (2005)	Drug court	Crim	Pub. Pol. J.
Labriola et al. (2005)	Batterer program	Crim	Think Tank
Goldkamp & White (2006)	Pretrial supervision	Crim	Crim. J.
Gottfredson et al. (2006)	Drug court	Crim	Crim. J.

(*Continued*)

Table 1 (*Continued*)

Reference	Intervention	Civ/Crim	Pub. Type
Labriola et al. (2008)	Batterer program	Crim	Spec. Law J.
Abrams & Rohlfs (2011)	Bail information	Crim	Econ. J.
Greiner & Pattanayak (2012)	Legal representation	Civ	Law Rev.
Greiner et al. (2012)	Legal representation	Civ	SSRN
Tomkins et al. (2012)	Court reminder	Crim	Spec. Law J.
Greiner et al. (2013)	Legal representation	Civ	Law Rev.
Rudd et al. (2015)	Parenting program	Civ	Psychol. J.

Fourth, a few of the studies we found were of reasonably high profile within the US legal profession (e.g., Freed & Wald 1964, Botein 1965, and Zeisel 1973 discussing Ayres et al. 1963; Neyfakh 2012 discussing Greiner & Pattanayak 2012). Thus, it is not the case that the US legal profession can field only low-profile RCTs.

Fifth, we found examples of complicated, multiyear, multisite, expensive, government- or foundation-funded RCTs in the law (e.g., Stapleton & Teitelbaum 1972). Thus, members of the US legal profession can mount, and have mounted, field operations bearing all of the accoutrements of the standard medical RCT [see Marks's (1997) description of the streptomycin/tuberculosis study].

Sixth, although we did not search exhaustively, what looking we were able to do generated no evidence that the results of an RCT in the US legal profession were actually used, in the sense that a program or policy changed because of the study's results. The closest we were able to find was that of Ayres et al. (1963) (see Zeisel 1973, pp. 111–12), which either contributed to or coincided with a movement to greater own-recognizance pretrial release in criminal cases (see Freed & Wald 1964). Thus, although "the absence of evidence is not evidence of absence,"[3] we are reminded that producing a rigorous evidentiary foundation for some part of the US legal profession's practices and persuading members of the profession to use that foundation are separate and formidable challenges.

Seventh, in the course of our search, we uncovered early (Zeisel et al. 1959) and periodic (e.g., Goldman 1983, Zeisel 1973) calls for the increased use of the randomized field experiment in the law. Thus, RCTs have remained scarce in the US legal profession despite advocacy on the subject. The reasons for the relative scarcity of RCTs in the US legal profession remain elusive.

THE UNITED STATES LEGAL PROFESSION'S RESISTANCE TO RCTS

In this section, we defend an assertion we have made previously, namely, that the US legal profession is, and has been, resistant to RCTs. One form of evidence to support our assertion would be numerous examples of well-designed, informative RCTs that researchers proposed to members of the US legal profession, but which were rejected. For obvious reasons, such evidence is difficult to find in the published literature. We are, however, fortunate enough to have had many such experiences ourselves. Studies we have proposed, and had rejected, include a follow-up to an already-completed study on the effectiveness of legal assistance (rejected on the grounds that randomizing legal assistance is unethical); proposals to evaluate early neutral evaluation and mediation programs in the federal courts (rejected on the grounds that no two cases are sufficiently alike

[3] The phrase is apparently from astronomer Martin Rees (see Morris 2014), not former Secretary of Defense Donald Rumsfeld.

for quantified study, even though they were all subject to the same ADR interventions); a design that would have assessed the effectiveness of limited assistance representation in a specialized state court (rejected with the comment that we do not need a randomized study to tell that light switches work); and an operation that would have assessed the effectiveness of differing content and format of letters urging indigent litigants to attend legal information clinics (we remain uncertain as to the basis for this rejection). Field studies often meet resistance, and ultimately do not go forward, for several reasons. Nonetheless, in each of these instances our experience was that resistance to the idea of randomizing the intervention played a role in the rejection of the study.

To this, we add that in the 80-odd years since the publication of the clinical trial by Amberson et al. (1931), US medical research has produced randomized field experiments in such volume that attempting a count would be a fool's errand. For new drugs and medical devices, the field RCT has become enshrined in law [Pub. Law No. 87–781, 76 Stat. 780, 781 (Oct. 10, 1962) (requiring "adequate and well-controlled studies"), see Harris 1981]. The US legal profession has not enshrined the RCT in the law. Instead, since the early 1930s, it has produced fewer than one randomized field experiment per year.

Finally, we supplement this discussion with three anecdotes. These anecdotes demonstrate that, even when researchers have been able to field RCTs in the US legal profession, lawyers and judges sometimes undermined them. And the lawyers and judges who did so appeared to have a common motive: certainty as to the "right" answer. We return to the theme of certainty in the final section.

The 1970s: CAMP and the Second Circuit

In the mid-1970s, attempting to respond to an influx of appeals, the Chief Judge of the United States Court of Appeals for the Second Circuit ordered lawyers in five pending appeals to attend settlement conferences over which he personally presided. Faced with encouragement to settle from the most powerful judge of the court that would eventually adjudicate their cases, and a potential member of the adjudicatory tribunal's panel, the parties in all five appeals reached agreement before oral argument.

The Chief Judge then constituted the Civil Appeals Management Program (CAMP). CAMP consisted of a mandatory one-hour preargument conference presided over by a lawyer working for the Second Circuit. The Chief Judge predicted that CAMP would reduce the rate at which cases reached the oral argument stage of litigation by 25 percentage points; he also set, as the key criterion for CAMP program effectiveness, an absolute minimum of a 10 percentage point reduction in the oral argument rate.

CAMP needed funding for its presiding lawyer. Funding was provided, but at the same time, the Federal Judicial Center designed and implemented an RCT. This first RCT showed no statistically significant reduction in oral argument rates in CAMP versus non-CAMP cases.

"The Second Circuit remained committed to the CAMP concept and maintained the program" (Patridge & Lind 1983, p. vii). A 1978 Second Circuit report credited CAMP as the reason why the Second Circuit, alone among the federal courts of appeals, closed more cases than it opened that year. And the Second Circuit requested that funding for CAMP continue.

The funding was provided, but a second RCT began. Before the second RCT could be finished, however, the Chief Judge found sufficient political clout to terminate it. The reason given: "the reluctance of the court to continue to exempt [control cases] from the CAMP program" (Patridge & Lind 1983, p. 2).

Prior to the Chief Judge's efforts, however, cases sufficient in number for a statistical analysis had made it through the second RCT's randomization protocol. The published analysis of those cases may be suspect because of an attempted adjustment for the fact that two different attorneys, rather

than a single attorney, conducted the CAMP conferences during the second RCT.[4] Accepting the results at face value, CAMP caused a 9.9 percentage point reduction in oral argument rates at $p = 0.049$. In other words, it failed the Chief Judge's prespecified absolute minimum success criterion. The Federal Judicial Center concluded, "the potential of [CAMP] is so great that all persons sharing responsibility for the management of appellate caseloads should give these procedures serious consideration" (Patridge & Lind 1983, p. vii).[5]

The 1990s: Mandatory Domestic Violence Counseling and Broward County

In the 1990s, interest arose in mandatory counseling for perpetrators of misdemeanor domestic violence. At the time, no randomized evaluation of a mandatory counseling program had been conducted. Researchers arranged one in Broward County, Florida, which already had a mandatory counseling program in place. Prior to the RCT's inception, the researchers initiated conversations with relevant stakeholders, including trial court judges and the prosecutor's office. The prosecutor's office opposed the study. With the support of the trial judges and the backing of the National Institute of Justice, the researchers persevered, fielding an RCT comparing one year of probation (control) to one year of probation plus six months of mandatory counseling (treated).

The prosecutor's office also persevered. It opposed the project on three fronts. First, it went to the press, encouraging reporters to ask questions comparing the experiment to those conducted by the Nazis. The press ultimately wrote balanced articles, and the RCT continued. Second, the prosecutor's office encouraged the probation office, victims' advocates, and other courthouse staff to create a difficult workplace for the RCT's research assistants. The efforts were successful in inducing high turnover among the research assistants and diverting the principal researchers' attention to constant training of new staff, but the RCT continued. Third, without giving notice to the researchers or trial judges, the prosecutor's office sought and obtained an appellate court order reversing trial court orders issued as part of the study. But the appellate court order was ambiguous. When the prosecutor's office filed an emergency motion seeking clarification and an immediate resentencing of all control group defendants, the appellate court reversed itself and held that it lacked jurisdiction. By this time, the field operation was largely complete.

The prosecutor's office contended throughout that an RCT was unethical and illegal because it denied victims whose batterers were in the control group the benefit of having their batterers undergo the mandatory counseling program. The RCT's results were as follows: Based on offender self-reports, victim reports, and official records, there was no statistically significant difference between treated and control groups in batterer rearrest rates, in (minor or severe) reabuse rates, or in batterer attitudes or beliefs regarding domestic violence (Feder & Dugan 2002).[6]

The 2010s: Representation in Agency Unemployment Appeals in Boston

Recently, one of us coauthored a study in partnership with a faculty-overseen, student-run clinic. The study evaluated the clinic's practice of representing claimants in internal agency appeals of claims adjuster decisions on unemployment benefits applications. The study randomized eligible claimants to either a clinic offer of representation (treated) or the provision of a list of other organizations that might provide assistance (control).

[4]This adjustment would be appropriate only if all CAMP conferences were to have been conducted by that single, specific attorney for their duration.

[5]The account above is taken from Goldman (1980), unless otherwise noted.

[6]The account above is taken from Feder et al. (2000).

As just suggested, at the time of the study, the student clinic was one of several organizations that represented claimants in these appeals. One such organization opposed the study from its outset, with some staff arguing that randomizing claimants to no offer of representation was unethical. When the researchers pointed out that all Boston area organizations providing claimant representation, including the one opposing the study, were oversubscribed (meaning that they regularly turned away eligible claimants requesting representation owing to lack of resources), the legal aid provider maintained its opposition.

When the study results became available, it turned out that approximately 39% of the control group had obtained representation (Greiner & Pattanayak 2012). Although far less than the treated group's representation rate (which was around 90%, id.), this figure was hard to understand given the aforementioned oversubscription. Subsequent conversations revealed that the legal services organization opposing the study, for the duration of the field operation, changed its intake practices. Before the study, it had considered any claimant seeking its assistance who had first been turned down by the student clinic (again, oversubscription) as equivalent to any other claimant. During the study, however, this organization prioritized representing study group control claimants. The organization later asserted that the study had caused it to have to stretch its resources.

The study found no statistically significant difference in claimant success rates between treated and control groups. It also found that the student clinic's offer of representation slowed adjudication by approximately three weeks, this in a setting in which—for those clients who are seeking benefits rather than defending an existing award—"timeliness is next to godliness" (Mashaw 1996, p. 19) because an eligible claimant ordinarily needs benefits as soon as possible. After the study's results became public, a different legal services provider contacted the research team, seeking assistance in conducting an RCT of its own unemployment representation practice. We were unable to help. In an effort to save resources, the provider that contacted us had earlier placed its intake system under the control of the organization that had opposed the study. The latter entity would not allow the former to conduct the RCT it desired.

WHY NOT LAW (AND WHY MEDICINE)? WHAT CAN BE DONE?

The previous sections make for grim reading. Except perhaps for the finding that RCTs in the US legal profession are (slightly) more numerous than we, at least, had anticipated, there is little to love. For those of us who think the present situation is bad, what can be done? We are not sure. But we are impatient for change, and that may mean exploring various mechanisms for inducing change despite uncertainty as to their effectiveness.

What mechanisms? Again, we are not sure. As noted above, however, we suggest that one way to generate ideas is to ask why and how US medicine chose to transform itself into a science, even as US law did not. Medicine is an apt foil for this analysis because it is another quintessential profession. In the United States, attorneys and physicians have traditionally shared four key attributes: expert knowledge, technical autonomy, an orientation toward the service of others, and high social status and income (Gorman & Sandefur 2011). Much has changed, for both professions. But the similarities between the two allow us to call into question some theories for why the RCT has failed to find a place in the US legal profession. And in our experience, legal professionals assert these theories as arguments for why RCTs are not compatible with judging or with the practice of law. Exposing these theories as myths, and digging deeper into the differences between law and medicine as they relate to RCTs, may prove fruitful.

One such theory/myth is that the RCT is incompatible with the concept of professional judgment. For example, judges have rejected our overtures regarding promising RCTs because they would temporarily have to give up a portion of their decision-making power, i.e., their professional

autonomy. A second such theory/myth is that each case or client is irreducibly complex. That is, no two cases (or clients) are sufficiently alike to allow one to think of them intelligently as elements of a portfolio, units susceptible to the application of one or more repeatable algorithms as opposed to requiring the application of individualized, expert knowledge.

Both arguments depend on the absolute primacy of professional judgment. But the past 80 years of experience with the RCT in medicine demonstrates that professional judgment and the RCT can coexist in various states of symbiosis. A historian of the first large-scale medical RCT, the investigation of streptomycin for the treatment of tuberculosis, states that what distinguished the study from prior research on the same drug "was not the involvement of specialists, but their apparent willingness in this instance to subordinate individual judgment to a common purpose" (Marks 1997, p. 126).[7] The RCT might be thought of as a tool to provide information to improve professional judgment. Or it might be thought of as a way to prevent professionals from pursuing ineffective, or economically unjustified, courses of action. Under the former view, RCT results improve the decisions of the professional, who maintains complete individual autonomy once the results are known. Under the latter view, RCT results restrain unproductive or destructive impulses of the (only human) professional, whose autonomy is circumscribed even after the results are known. Either way, even if one is not inclined to question the premise that leaving critical decisions in the hands of human professionals is always best (we are so inclined), the past 80 years in medicine show that although there may be friction between the RCT and professional judgment, there is no fundamental incompatibility.

Another theory/myth of why RCTs are incompatible with judging or with the practice of law focuses on professional ethics. For example, in a response to an early version of a paper one of us coauthored, an attorney contended that if a legal services provider used an RCT to evaluate a program providing full representation to eligible clients, its attorneys would be required to advocate for clients with frivolous cases, violating their duty of candor to the tribunal (Monsell 2011). And scholars have explored the tension between a professional's duty to advance her client's or patient's individual best interests and her duty to accumulate knowledge to further the best interests of future patients (Fried 1974).

Once again, the medical example demonstrates that any tension between RCTs and professional ethics is resolvable. Regarding the duty of candor to the tribunal, as Greiner (2011) explained at the time, there need be no conflict. A merits screen applied at intake, and a motion to withdraw from representation when factual investigation shows a lack of merit, can be defined respectively as a study eligibility criterion and a part of the service being evaluated. In medicine, research physicians impose all sorts of eligibility criteria in drug trials, and a doctor might retain the option to remove clinical trial patients who adopted a post-enrollment course of action incompatible with the study (e.g., resuming a smoking habit inimical to a drug's active ingredient).

The accounts above are cursory and incomplete. One could argue each human body and its ailments are not irreducibly complex, but each legal case or client is. We doubt that to be the case. Modern complex litigation (class actions, multidistrict litigation) consists in part of aggregating claims into portfolios of similar units [e.g., Federal Rule of Civil Procedure 23(a)'s commonality requirement], and at the other end of the scale, debt buyers have commoditized collection litigation (Jiménez 2015). And even if legal, but not medical, cases now appear irreducibly complex, perhaps that is because we are viewing matters on the right side of eight decades of medical RCTs. But

[7] To reiterate, our point here is not that all is rosy in the medical profession, just that the situation in medicine is vastly superior to that in law.

because at a broad level professional judgment and professional ethics are common to both law and medicine, neither is likely to be a fruitful focus for exploration as to our questions of interest.

So what would be a fruitful focus for exploration? For the third time, we are unsure. But to close this paper, we offer some speculation. Our speculation is built on the assumption that the desire to deploy RCTs must stem from an internalization of uncertainty. Without a deep sense of uncertainty about how the world works, there is little reason to seek methodologies for producing knowledge external to one's own experiences and those of elites within one's profession. And as at least one scholar would have it, medicine's adoption of the RCT stemmed in part from a desire among elite physicians to establish an irrefutable methodology to distinguish effective therapies from commercially promoted snake oil (Marks 1997). Such a project had to depend on the elite physicians' ability to persuade their contemporaries to view the claims of commercial interests, and the knowledge gained by their own experiences, with skepticism.

Are there distinctions (we speculate socially constructed ones) between law and medicine that might lead present-day practitioners of one but not the other to internalize a norm of acknowledging uncertainty? We are not the right scholars to address this question. We lack training in history, history of science, sociology, epistemology, psychology, and most other fields directly relevant to this inquiry. At present, all we can offer are two thoughts based, alas, on our own experiences.

Our first thought concerns the nature of the two professions' diagnostic thought processes. The stereotypical physician confronted with an ailing patient begins an inquisitive diagnostic process. The physician's diagnostic process is supposed to address the question, "What is wrong?" The physician does not initially know, and she searches for an answer. For the stereotypical lawyer, however, the process of diagnosis is less inquisitive than instrumental. The client states her desires (perhaps with varying degrees of clarity), in the form of the status she wants to achieve (e.g., no longer married) or the goal she wishes to reach (e.g., avoiding or minimizing imprisonment, obtaining benefits, remaining in an apartment). Within wide boundaries, the stereotypical lawyer's diagnostic process is not a search for answers. Answers come from the client. The stereotypical lawyer's diagnostic process is a search for arguments. At least for litigation or quasi-adversarial settings, analysis is in the service of advocacy from square one. Thus, our speculation is that because lawyers train themselves in instrumental, as opposed to inquisitive, analysis, the result is ingrained arrogance. There is less willingness to acknowledge the kind of uncertainty needed to make an RCT relevant.

Our second strand of speculation is related but easier to state: Perhaps there is an implicit belief among lawyers and judges that to appear certain one must be certain, and that there is social value in appearing certain. As present and former litigators, we (the authors) believe that to be maximally effective advocates, we must temporarily believe in the rightness of the positions that we argue. Similarly, judges might doubt that there would be as much voluntary compliance with law if, each time a judge sentenced a criminal defendant to a period of incarceration, the judge stated on the record, "I do not know if confining you will do you any good. I also do not know if confining you will deter others, or give the victims of your crime closure, or will otherwise serve any useful purpose. But I will nevertheless sentence you to x years." Instead, some judges may have internalized the belief that they must appear certain to be effective, and to appear certain they must be certain.

One might argue that the same could be true in medicine. Perhaps physicians believe that they must persuade patients that they are certain about a proposed treatment for that treatment to be effective, and to persuade patients of certainty physicians must actually be certain. We doubt it. Our own physicians do not behave this way. The fact that ideal drug trials are double-blind suggests that the medical community does not uniformly think this way. And one might argue that it is harder to tell a story of how a physician's certainty with an individual patient during

a presumptively confidential consultation has as much of an effect on society as a judge's public sentencing of a criminal defendant has on other citizens. Our speculation is that the norm of certainty is stronger in lawyers than in physicians.

For the umpteenth time, we emphasize that all of this is speculation. There are other hypotheses that are worth testing, perhaps with behaviorally or psychologically informed survey research. But whatever portion of the above is true would appear to us to be socially constructed and, perhaps, subject to change. For those of us who care about transforming law into a field less hostile to the RCT, that may be cause for hope.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors thank Ellen Degnan, Don Green, Kevin Quinn, Chris Robertson, Ameet Sarpatwari, Roseanna Sommers, Dane Thorley, and David Yokum for helpful comments. All mistakes are our own.

LITERATURE CITED

Abramowicz M, Ayres I, Listokin Y. 2011. Randomizing law. *Univ. Penn. Law Rev.* 159:929

Abrams DS, Rohlfs C. 2011. Optimal bail and the value of freedom: evidence from the Philadelphia Bail Experiment. *Econ. Inq.* 49(3):750–70

Abrams DS, Yoon AH. 2007. The luck of the draw: using random case assignment to investigate attorney ability. *Univ. Chic. Law Rev.* 74(4):1145–77

Albiston CR, Sandefur RL. 2013. Expanding the empirical study of access to justice. *Wis. Law Rev.* 2013:101–20

Amberson JB, McMahon BT, Pinner M. 1931. A clinical trial of sanocrysin in pulmonary tuberculosis. *Am. Rev. Tuberc.* 24:401–35

Anderson JM, Heaton P. 2012. How much difference does the lawyer make? The effect of defense counsel on murder case outcomes. *Yale Law J.* 122:154–217

Ayres CE, Rankin A, Sturz H. 1963. The Manhattan Bail Project: an interim report on the use of pretrial parole. *N.Y. Univ. Law Rev.* 38:67

Baker SH, Sadd S. 1981. *Diversion of Felony Arrests: An Experiment in Pretrial Intervention*. Washington, DC: US Dep. Justice, Natl. Inst. Justice

Berecochea JE, Jaman DR, Jones WA. 1973. *Time served in prison and parole outcome: an experimental study.* Tech. Rep. 1, Calif. Dep. Correct. Res. Unit, Sacramento

Boissel JP. 1989. Impact of randomized clinical trials on medical practices. *Control. Clin. Trials* 10(4):120S–34S

Botein B. 1965. The Manhattan Bail Project: its impact on criminology and the criminal law process. *Texas Law Rev.* 43:319

Chilton AS, Levy MK. 2015. Challenging the randomness of panel assignment in the Federal Courts of Appeals. *Cornell Law Rev.* 101:1–55

Clark SH, Ellen ED, McCormick K. 1995. *Court-ordered civil case mediation in North Carolina: an evaluation of its effects.* Tech. Rep., N.C. Adm. Off. Courts, Chapel Hill. **http://courtadr.org/files/CourtOrderedMedNC.pdf**

Cohen IG. 2013. Rationing legal services. *J. Legal Anal.* 5:221

Colbert DL, Paternoster R, Bushway S. 2002. Do attorneys really matter? The empirical and legal case for the right of counsel at bail. *Cardozo Law Rev.* 23(5):1719–93

Davis RC, Taylor BG, Maxwell CD. 2000. *Does batterer treatment reduce violence? A randomized experiment in Brooklyn*. Tech. Rep. 180772, Urban Inst., Washington, DC. **https://www.ncjrs.gov/pdffiles1/nij/grants/180772.pdf**

Deschenes EP, Turner S, Greenwood PW. 1995. Drug court or probation?: An experimental evaluation of Maricopa County's drug court. *Justice Syst. J.* 18(1):55–73

Ditman KS, Crawford GC, Forgy EW, Moskowitz H, MacAndrew C. 1967. A controlled experiment on the use of court probation for drunk arrests. *Am. J. Psychiatry* 124(2):160–63

Eaglin JB. 1990. *The pre-argument conference program in the Sixth Circuit Court of Appeals: an evaluation*. Tech. Rep. FJC-R-90-1, Fed. Judicial Cent., Washington, DC. **http://www.fjc.gov/public/pdf.nsf/lookup/preargcon.pdf/$file/preargcon.pdf**

Elek JK, Ware LJ, Ratcliff JJ. 2012. Knowing when the camera lies: judicial instructions mitigate the camera perspective bias. *Legal Crim. Psychol.* 17:123–35

Engel C. 2013. Legal experiments—Mission impossible? **http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2276566**

Farrington DP. 1983. Randomized experiments on crime and justice. In *Crime and Justice: A Review of Research*, Vol. 4. Chicago: Univ. Chicago Press

Farrington DP, Welsh BC. 2005. Randomized experiments in criminology: What have we learned in the last two decades? *J. Exp. Criminol.* 1:9–38

Feder L, Dugan L. 2002. A test of the efficacy of court-mandated counseling for domestic violence offenders: the Broward experiment. *Justice Q.* 19(2):343–75

Feder L, Jolin A, Feyerherm W. 2000. Lessons from two randomized experiments in criminal justice settings. *Crime Delinq.* 46:380–400

Finkelstein A, Taubman S. 2015. Randomized evaluations to improve health care delivery. *Science* 13:720–22

Freed DJ, Wald PM. 1964. *Bail in the United States*. Washington, DC: US Dep. Justice, Vera Inst.

Fried C. 1974. *Medical Experimentation*. Amsterdam: North-Holland Publ. Co.

Gaudet FJ, Harris GS, St. John CW. 1933. Individual differences in the sentencing tendencies of judges. *J. Crim. Law Criminol.* 23(5):811–18

Goldkamp JS, Goddfredson MR. 1984. *Judicial Guidelines for Bail: The Philadelphia Experiment*. Washington, DC: US Dep. Justice, Natl. Inst. Justice

Goldkamp JS, White MD. 2006. Restoring accountability in pretrial release: the Philadelphia pretrial release supervision experiments. *J. Exp. Criminol.* 2:143–81

Goldman J. 1978. The Civil Appeals Management Plan: an experiment in appellate procedural reform. *Columbia Law Rev.* 78(6):1209–40

Goldman J. 1980. *Ineffective Justice: Evaluating the Preappeal Conference*. Thousand Oaks, CA: Sage

Goldman J. 1983. Experimentation in the law: report of the Federal Judicial Center Advisory Committee on Experimentation in the Law. *Am. Bar Found. Res. J.* 8(3):733–39

Gorman EH, Sandefur RL. 2011. 'Golden age,' quiescence, and revival: how the sociology of professions became the study of knowledge-based work. *Work Occup.* 38(3):275–302

Gottfredson DC, Exum ML. 2002. The Baltimore City Drug Treatment Court: one-year results from a randomized study. *J. Res. Crime Delinq.* 39(3):337–56

Gottfredson DC, Kearley BW, Najaka SS, Rocha CM. 2005. The Baltimore City Drug Treatment Court 3-year self-report outcome study. *Eval. Rev.* 29(1):42–64

Gottfredson DC, Najaka SS, Kearley B. 2003. Effectiveness of drug treatment courts: evidence from a randomized trial. *Criminol. Public Policy* 2(2):171–96

Gottfredson DC, Najaka SS, Kearley BW, Rocha CM. 2006. Long-term effects of participation in the Baltimore City Drug Treatment Court: results from an experimental study. *J. Exp. Criminol.* 2(1):67–98

Green DP, Thorley DR. 2014. Field experimentation and the study of law and policy. *Annu. Rev. Law Soc. Sci.* 10:53–72

Greenwood PW, Turner S. 1993. Evaluation of the Paint Creek Youth Center: a residential program for serious delinquents. *Criminology* 31(2):263–79

Greiner DJ, Pattanayak CW. 2012. Randomized evaluation in legal assistance: What difference does representation (offer and actual use) make? *Yale Law J.* 121(8):2118–215

Greiner DJ, Pattanayak CW, Hennessy JP. 2012. How effective are limited legal assistance programs? A randomized experiment in a Massachusetts housing court. Accessed Sept. 13, 2015. **http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1880078**

Greiner DJ, Pattanayak CW, Hennessy JP. 2013. The limits of unbundled legal assistance: a randomized study in a Massachusetts district court and prospects for the future. *Harv. Law Rev.* 126(4):901–89

Greiner J. 2011. Randomized trials do not require abandoning merits screens or ethical duties: a response to Margaret Monsell. *Concurring Opinions*, March 28. **http://concurringopinions.com/archives/2011/03/randomized-trials-do-not-require-abandoning-merits-screens-or-ethical-duties-a-response-to-margaret-monsell.html**

Hannaford PL, Hans VP, Munsterman GT. 2000. Permitting jury discussions during trial: impact of the Arizona reform. *Law Hum. Behav.* 24(3):359–82

Hanson R, Hersey GW. 1991. Appellate court congestion, or how do you spell R-E-L-I-E-F? *Gov. Fla.* 1:11

Harrell A, Cavanagh S, Roman J. 2000. *Final report: findings from the evaluation of the D.C. Superior Court Drug Intervention Program*. Tech. Rep., Urban Inst., Washington, DC. **https://www.ncjrs.gov/pdffiles1/nij/178941.pdf**

Harris R. 1981. *The Real Voice*. Upper Saddle River, NJ: Regents/Prentice Hall

Heuer L, Penrod SD. 1988. Increasing jurors' participation in trials: a field experiment with jury notetaking and question asking. *Law Hum. Behav.* 12(3):231–61

Heuer L, Penrod SD. 1989. Instructing jurors: a field experiment with written and preliminary instructions. *Law Hum. Behav.* 13(4):409–30

Heuer L, Penrod SD. 1994a. Juror notetaking and question asking during trials: a national field experiment. *Law Hum. Behav.* 18(2):121–50

Heuer L, Penrod SD. 1994b. Trial complexity: a field investigation of its meaning and its effects. *Law Hum. Behav.* 18(1):29–51

Imbens GW, Rubin DB. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge, UK: Cambridge Univ. Press

Jiménez D. 2015. Dirty debts sold dirt cheap. *Harv. J. Legis.* 52:41

Kakalik JS, Dunworth T, Hill LA, McCaffrey D, Oshrio M, et al. 1996. *An evaluation of mediation and early neutral evaluation under the Civil Justice Reform Act*. Tech. Rep., Inst. Civ. Justice, Santa Monica, CA. **http://www.rand.org/content/dam/rand/pubs/monograph_reports/2007/MR803.pdf**

Kobbervig W. 1991. *Mediation of civil cases in Hennepin County: an evaluation*. Tech. Rep., Minn. Judic. Cent. Res. Plan. Off. State Court Adm., St. Paul, MN. **http://archive.leg.state.mn.us/docs/pre2003/mandated/910535.pdf**

Labriola M, Rempel M, Davis RC. 2005. *Testing the effectiveness of batterer programs and judicial monitoring*. Tech. Rep., Cent. Court Innov., New York. **http://www.courtinnovation.org/sites/default/files/battererprogramseffectiveness.pdf**

Labriola M, Rempel M, Davis RC. 2008. Do batterer programs reduce recidivism? Results from a randomized trial in the Bronx. *Justice Q.* 25(2):252–82

Mandell MB, Marshall A. 2002. *The effects of court-ordered mediation in workers' compensation cases filed in circuit court: results from an experiment conducted in the circuit court for Baltimore City*. Tech. Rep., Univ. Md. Baltim. Cty. Md. Inst. Policy Anal. Res., Baltimore. **http://www.courts.state.md.us/macro/pdfs/reports/baltcityworkercompreportfinal.pdf**

Margolin FM. 1970. Counseling in contested custody and visiting in a court setting. *Family Court Rev.* 8(2):11–13

Marks HM. 1997. *The Progress of Experiment: Science And Therapeutic Reform in the United States, 1900–1990*. Cambridge, UK: Cambridge Univ. Press

Marlowe DB, Festinger DS, Lee PA, Schepise MM, Hazzard JE, et al. 2003. Are judicial status hearings a key component of drug court? During-treatment data from a randomized trial. *Crim. Justice Behav.* 30(2):141–62

Mashaw J. 1996. Unemployment compensation: continuity, change, and the prospects for reform. *Univ. Mich. J. Law Reform* 29:1–24

McEwen CA. 1992. *An evaluation of the ADR pilot project*. Final Rep., Bowdoin Coll. Accessed Sept. 13, 2015. **http://www.bowdoin.edu/faculty/c/cmcewen/pdfs/an-evaluation-of-the-adr-pilot-project-final-report-1992.pdf**

McGarrell EF, Olivares K, Crawford K, Kroovand N. 2000. *Returning justice to the community: the Indianapolis Juvenile Restorative Justice Experiment*. Tech. Rep., Hudson Inst. Crime Control Policy Cent., Indianapolis. **http://www.ibarji.org/docs/mcgarrell.pdf**

Meldrum ML. 2000. A brief history of the randomized control trial: from oranges and lemons to the gold standard. *Hematol./Oncol. Clin. N. Am.* 14(4):745–60

Monsell M. 2011. What difference representation: case selection and professional responsibility. *Concurring Opinions*, March 28. **http://concurringopinions.com/archives/2011/03/what-difference-representation-case-selection-and-professional-responsibility.html**

Morris E. 2014. The Certainty of Donald Rumsfeld (Part 4): Absence of evidence isn't evidence of absence. *The New York Times*, March 28. **http://opinionator.blogs.nytimes.com/2014/03/28/the-certainty-of-donald-rumsfeld-part-4/#ftn3**

Neyfakh L. 2012. Free legal advice—But does it work? *The Boston Globe*, Dec. 23, **https://www.bostonglobe.com/ideas/2012/12/23/free-legal-advice-but-does-work/GUoKaBqK068IhaZt5Bj-qKI/story.html**

Parker M. 2005. False dichotomies: EBM, clinical freedom, and the art of medicine. *Med. Humanit.* 31:23–30

Patridge A, Lind A. 1983. *A Reevaluation of the Civil Appeals Management Plan*. Washington, DC: Fed. Judic. Cent.

Pearson J, Thoennes N. 1984. Mediating and litigating custody disputes: a longitudinal evaluation. *Fam. Law Q.* 17(4):497–524

Petersilia J, Turner S. 1991. An evaluation of intensive probation in California. *J. Crim. Law Criminol.* 82(3):610–58

Powell v. Alabama, 287 U.S. 45, 69 (1932)

Robinson CT, Yokum DV. 2012. The effect of blinded experts on juror verdicts. *J. Empir. Legal Stud.* 9:765

Rosenberg M. 1964. *The Pretrial Conference and Effective Judgment*. New York: Columbia Univ. Press

Rudd BN, Hotzworth-Munroe A, Reyome JG, Applegate AG, D'Nofrio BM. 2015. Randomized control trial: online parent program and waiting period for unmarried parents in Title IV-D court. *J. Fam. Psychol.* 29(5):679–85

Samaha AM. 2009. Randomization and adjudication. *William Mary Law Rev.* 51:1–86

Schneider AL. 1986. Restitution and recidivism rates of juvenile offenders: results from four experimental studies. *Criminology* 24(3):533–52

Seron C, Frankel M, Van Ryzin G, Kovath J. 2001. The impact of legal counsel on outcomes for poor tenants in New York City's housing court: results of a randomized experiment. *Law Soc. Rev.* 35(2):419–34

Stapleton WV, Teitelbaum LE. 1972. *In Defense of Youth: A Study of the Role of Counsel in American Juvenile Courts*. New York: Russell Sage Found.

Steelman DC, Goldman J, Conti SD. 1986. *The Settlement Conference: Experimenting with Appellate Justice: Final Report*. Williamsburg, VA: Natl. Cent. State Courts

Susskind R. 2010. *The End of Lawyers? Rethinking the Nature of Legal Services*. Oxford: Oxford Univ. Press

Thoennes N. 2001. *Permanent custody mediation*. Tech. Rep., Cent. Policy Res., Denver, CO. **http://cntrpolres.qwestoffice.net/reports/Lucas%20County.pdf**

Thorley D. 2015. *Randomness Pre-Considered: Recognizing and Accounting for "De-Randomizing" Events When Utilizing Random Judicial Assignment*. Nov. 20. Unpublished manuscript, copy on file with authors

Tomkins AJ, Bornstein B, Herian MN, Rosenbaum DI, Neeley EM. 2012. An experiment in the law: studying a technique to reduce failure to appear in court. *Court Rev.* 48:96–106

Ursa Inst. 1984. *Evaluation of the Early Representation by Defense Counsel Field Test*. Washington, DC: US Dep. Justice, Natl. Inst. Justice

Weitzman SE. 1983. *Early Representation by Defense Counsel in Shelby County, Tennessee 1982–1983: Final Report*. Washington, DC: US Dep. Justice, Natl. Inst. Justice
Zeisel H. 1973. Reflections on experimental techniques in the law. *J. Legal Stud.* 2(1):107–24
Zeisel H, Kalven H Jr, Buchholz B. 1959. The case for the official experiment. In *Delay in the Court*, ed. H Zeisel, Chapter 21. Boston: Little, Brown
Zuberbuhler J. 2001. Early intervention mediation: the use of court-ordered mediation in the initial stages of divorce litigation to resolve parenting issues. *Fam. Court Rev.* 39:203–5