

Causal Stories in *Vergara v. California*

Educational Policy

1–35

© The Author(s) 2018

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0895904818802110

journals.sagepub.com/home/epx



Jessica J. Gottlieb¹, Ethan L. Hutt² ,
and Benjamin M. Superfine³

Abstract

In 2012, families in California filed a lawsuit alleging that five state statutes governing teacher tenure, dismissal, and seniority together violate the state constitution's requirements for equal protection. Central to the case were competing narratives about the relationship between these statutes, the work of teachers, and the achievement of students. This article analyzes those narratives utilizing the trial court transcripts and judicial opinions in *Vergara v. California*. We find that despite reaching divergent rulings, the trial and appellate courts provided highly typified accounts of the case—ones that emphasized individual agency and dismissed or deemphasized the importance of the social and political context of schooling. These findings are important for understanding how complex policy debates become transformed within legal proceedings and for understanding the capacity of courts to engage complex evidence and narratives—a major issue given that courts remain an important venue for school reform.

Keywords

accountability, teacher qualifications, evaluation and assessment, legal issues, education reform, causal inference

¹Texas Tech University, Lubbock, USA

²University of Maryland, College Park, USA

³The University of Illinois at Chicago, USA

Corresponding Author:

Ethan L. Hutt, Assistant Professor, 2203 Benjamin Building University of Maryland, College Park, MD 20742-1175, USA.

Email: ehutt@umd.edu

Introduction

In May 2012, nine students and their families filed a lawsuit in Los Angeles Superior Court alleging that five state statutes governing teacher tenure, dismissal, and seniority, when enforced together, violate the state constitution's requirements for equal protection. According to the plaintiffs, the combined effect of these laws produces a "vicious cycle" in which ineffective teachers receive tenure, become difficult to remove, and remain protected by seniority rules even in times of teacher layoffs (*Vergara v. California*, 2014, p. 16). The constraints these laws place on administrator discretion, combined with the high turnover in low-performing schools, result, according to the plaintiffs, in systematically burdening poor and minority students with the least effective—"grossly ineffective"—teachers. This lawsuit, *Vergara v. California* (2014), at once highlighting the noncontroversial point that teachers are crucially important to quality education and the controversial point that tenure laws are an obstacle to obtaining these results in practice, received national attention and was watched carefully by commentators on all sides (e.g., Boutros, 2014; Layton, 2014). Of particular interest was the novel legal claim advanced by the plaintiffs—that teacher quality was a component of equitable education—and the possibility that, if successful, the case could inspire similar lawsuits in other states (A. Baker, 2014; e.g., *Forslund v. Minnesota*, 2016), with potentially wide-ranging impact on teacher tenure laws (Sawchuk, 2014).

Given the perceived stakes of the case—"The War on Teacher Tenure," read the *Time* cover story headline (Edwards, 2014)—much of the coverage framed the case either as a proxy fight against teachers' unions and the merits of tenure or as a referendum on the merits of "value-added models"—the basis for the plaintiffs' theory about unequal access to quality teachers. These issues concerning teacher policy have, likewise, received an increasing amount of scholarly attention as research has traced efforts to curb the influence of teachers' unions (e.g., Giersch, 2014; McGuinn, 2012, Superfine & Gottlieb, 2014), assess the merits of value-added models (e.g., D. Harris, 2011; National Research Council, 2010), and understand the legal implications of their adoption (e.g., Green, Baker, & Oluwole, 2012; Hutt & Tang, 2013; Pullin, 2013). While these debates swirled outside the courtroom, less attention has been given to how these lines of research and the policy arguments they implied were presented, reconstructed, and ruled on *inside* the courtroom. This lack of attention to how these arguments unfolded in courtrooms constitutes a considerable oversight, given, first, that judicial determinations can have profound effects on which education policies are available for legislators and school officials to enact and, second, that, given this policy

influence, scholars have long recognized that judicial proceedings differ in important ways from other venues of deliberation and decision making because of their formal rules, notions of expertise, standards of evidence, and determinations of harm (e.g., Ewick, Kagan, & Sarat, 1999; Kagan, 2003; Komesar, 1994). Examining, then, how judges conceive of the operation of schools and weigh empirical evidence about schools is an important part of our understanding of education policy, in general, and teacher policy, in particular. This is especially true at a time when reformers across the ideological spectrum are turning to courts to achieve education policy and education reform victories (Superfine & Thompson, 2016).

In this article, we address this gap by examining the ways in which narratives about teachers and teaching were constructed during the *Vergara* trial and, subsequently, taken up in the trial court and appellate court decisions. Specifically, we use the trial court transcripts and judicial opinions to examine how the narratives presented in *Vergara* consider the issue of teacher quality, evaluation, and supply and of how these depictions incorporated (or excluded) the broader organizational, political, social, and economic contexts in which teachers' work is situated. Attention to these narratives is particularly important given that the plaintiffs sought to link a set of statutes concerning teacher personnel decisions to the provision of effective teaching and, ultimately, to educational attainment. Establishing (or disputing) such links in a court of law requires the crafting causal stories that link each of these elements in a chain. By examining this chain and the theories of teaching and school operation from which they are forged, we can arrive at a clearer understanding of how narratives are constructed and selected in the course of a trial. To the extent that different narratives reflect different theories about the work, the work environment, and the agency of teachers, their embrace by courts in cases like *Vergara* can have important policy implications.

It is important to say at the outset that in analyzing the narratives presented at trial and constructed in the subsequent court rulings, we are in no way suggesting that these are the only available narratives one could tell about teachers or about the general operation of schools. In the course of litigation, the parties must make strategic decisions about what evidence to present and which stories to tell based on a variety of factors, including the specific laws in question and the relevant legal precedents. Similarly, even those who are skeptical that judges are hemmed in by "The Law" recognize that the presiding judges are constrained by the stated claim, the evidence presented at trial, and the controlling precedents. That these idiosyncrasies might have important bearings on how cases are decided only underscores the importance of examining these issues.

In conducting this analysis, we find, consistent with the literature in legal studies literature, that the narratives constructed by both the trial court and the appellate court provided highly “typified accounts” (Lempert & Sanders, 1986) of the operation of California’s schools. That is, both the trial court and appellate court judges engaged in storytelling that emphasized agency—the teachers’ and the administrators’, respectively—and that minimized the extent to which they considered broader social context. This was true despite the fact that the trial court and appellate court came to divergent conclusions about the *legal* merits of the plaintiffs’ case. Given that *Vergara* represents the intermingling of a decades-long federal and state policy trend to increase teacher evaluation and accountability (e.g., Cochran-Smith, 2001; National Council on Teacher Quality, 2013; Superfine, Gottlieb, & Smylie, 2012) and a long tradition of using courts to pursue educational equality (e.g., Rebell & Block, 1982; Ryan, 2007; Wise, 1968), it is important that we analyze the narratives told about teachers and schools in these settings not only for the advocates on each side of these debates but also for scholars whose work might be used in this contexts and for those interested in the capacities of courts to engage in complex story-telling and challenge of school reform.

Policy and Legal Background

The *Vergara* case reflects the confluence of two policy trends: one aimed at linking teacher evaluation more tightly to measures of student learning growth; the other aimed at reducing the influence of teacher unions and collective bargaining by tying personnel decisions to measures of performance. In challenging the constitutionality of California’s statutes governing teacher tenure, dismissal, and seniority, *Vergara* reflects an attempt to advance these policy trends not just in the legislative arena by changing the laws but rather to advance them through the judicial arena by having the existing laws overturned and establishing a new legal determination on the relationship between tenure laws and the provision of educational equality.

Researchers and policy makers have come to a consensus that teachers are one of the most important in-school factors influencing students’ learning opportunities and performance (Darling-Hammond, 2000; Goe, 2007; McGuinn, 2010). In light of this consensus, policies focused on teacher evaluation and accountability have recently gone through dramatic modifications and expansions. While states have long held legal authority over the basic requirements for teacher quality, such as those governing licensure, teacher education program approval, professional development, and many aspects of compensation (Corcoran, 2007), teachers generally have not received robust feedback for improvement and/or exhortations to improve on any consistent

basis (Glazerman et al., 2011). Indeed, teacher evaluation systems have generally failed at distinguishing between effective and ineffective teachers (Kennedy, 2010).

While states have been primarily responsible for the development and implementation of teacher evaluation and accountability systems, the federal government has become increasingly involved in teacher workforce policy (Superfine et al., 2012). For example, with the highly qualified teacher mandate in No Child Left Behind (NCLB), the U.S. Congress highlighted teacher quality as a key object of federal policy. The Race to the Top Fund (RTT), contained in the American Recovery and Reinvestment Act of 2009 (Pub. L. No. 111-5, § 14005-06), included financial incentives for states to enact accountability-based teacher evaluation systems, and waivers releasing states from certain requirements under NCLB incentivized such changes as well (Superfine et al., 2012). Partially in response to this increased federal involvement, state policy makers recently have focused on intensifying teacher evaluation and holding teachers accountable for their performance under the belief that such an approach can increase the quality and effectiveness of teachers. Since the late 2000s, a majority of states have passed laws requiring school districts to develop systems for evaluating and holding accountable teachers for their performance, partly on the basis of student achievement growth (National Council on Teacher Quality, 2013). For example, while only 15 states required teacher evaluations to include “objective measures” of student achievement in 2009, 41 states required such evaluations by 2013 (National Council on Teacher Quality, 2013). While the speed of implementation of these new evaluation systems has varied by state and has not been without setbacks, the goal of tying teacher personnel decisions to student achievement has become a national policy trend.¹

Even as these evaluation models have been adopted nationwide, they have been the source of considerable debate among education researchers. Many scholars argue that value-added models represent a substantial increase in the quality, rigor, and reliability of teacher evaluation efforts, especially when compared with traditional proxies for teacher quality, such as years of experience or traditional assessment tools like in-class observations (e.g., Hanushek & Rivkin, 2012; Kane, Rockoff, & Staiger, 2008). Proponents argue, additionally, that traditional evaluation systems fail to discern degrees of effectiveness or even between effective and ineffective teachers as many systems rate nearly all of their teachers as effective (e.g., Weisberg et al., 2009). Other scholars, however, have raised questions about the possible technical shortcomings, including bias resulting from nonrandom assignment of students to teachers, the instability of the metrics when based on limited years of data, and the sensitivity of the measures to the model specification (e.g., American

Educational Research Association, 2015; Baker, Oluwole, & Green, 2013; National Research Council, 2010). These potential issues, they argue, call into question the validity and the wisdom of decisions made on the basis of these measures. Irrespective of these issues over accuracy or validity, other researchers have raised questions about the “uncertain” theory of action underlying these evaluation systems—specifically whether these systems are capable of boosting teacher motivation or instructional abilities (Kennedy, 2010). Although these debates remain unsettled, teacher accountability systems continue to be a core strategy employed by states to improve teacher effectiveness.

Concurrent with these changes in teacher evaluation and accountability policies, many states made alterations to their policies governing public sector collective bargaining, with teachers’ unions in particular facing significant political challenges along multiple fronts (Malin, 2012). Since 2010, at least 12 states have modified their general collective bargaining laws to narrow the range of issues over which public sector workers can bargain (Malin, 2012). In education, these modifications have included lengthening the time it takes for teachers to achieve tenure, streamlining teacher discipline and removal procedures, and tying teacher pay to evaluation (Koski, 2012). Such changes are largely aligned with new teacher evaluation requirements in states that often remove issues related to teacher evaluation from the collective bargaining process. For instance, Wisconsin’s Act 10 restricted the issues that teachers could collectively bargain to “base wages” and barred collective bargaining over such traditional issues as benefits, hours, and working conditions (Malin, 2012). High-profile legal attacks on public sector unions have begun to emerge as well. In *Harris v. Quinn* (2014), the U.S. Supreme Court ruled that a state law allowing unions to collect fees from certain types of nonunion employees was unconstitutional and subsequently ruled in *Janus v. American Federation of State, County, and Municipal Employees* (2018) that the collection of agency fees from non-consenting public-sector employees is unconstitutional.

Beyond their use in contract negotiations and personnel decisions, teacher evaluation systems have also begun to emerge in the courts as a critical element of cases aimed at promoting broader education reform. Courts have long been a key institution for promoting such reform, primarily through public law litigation (e.g., Chayes, 1976; Cole, 2016). In this type of litigation, courts generally craft broad, forward-looking remedies aimed at reforming institutional structures to address systemic social problems. Indeed, the courts have engaged in several high-profile public law cases in education, such as in desegregation and school finance reform (Superfine, 2013). In *Martinez v. State of New Mexico* (2014), the Mexican American Legal

Defense Fund (MALDEF) challenged the constitutionality of a state's teacher evaluation system because it failed to provide for effective monitoring and improvement of teachers, ultimately leaving many students without experienced and well-trained teachers. While MALDEF also cited many other problems with the state's education system, the alleged problems inherent in the state's teacher evaluation system constituted a central target.

Teacher evaluations systems have also emerged in the courts as teachers have looked to the legal process for protection against what they have perceived to be the undesirable and unfair effects of teacher evaluation systems. Teachers who feel that they have been treated unfairly by the implementation of teacher evaluation systems have filed several lawsuits (Pullin, 2013). In *Cook v. Stewart* (2014), teacher challenged Florida's teacher evaluation policy and argued that the policy was irrational and arbitrary. The teacher pointed out, in particular, that some teachers were evaluated on the basis of test scores earned by students whose teachers teach in subjects not associated with tests, and that some teachers are evaluated using the scores of students that these teachers do not teach at all. In *Urbanski v. King* (2014), the Rochester Teachers Association filed a lawsuit arguing that New York's teacher evaluation policy violated teachers' rights to fair evaluations because it failed to account for the effects of poverty on student achievement, unfairly penalizing Rochester teachers. Indeed, the types of arguments made against and in favor of teacher evaluation policies in the courts appear to parallel the arguments in the broader policy sphere.

Vergara reflects the convergence of recent policy and legal trends focused on teacher evaluation and performance and the attending causal stories that tie teacher actions to value-added models. As with these earlier cases, *Vergara* required judges to evaluate the plausibility of the basic causal stories embedded in the state's evaluation policy and conveyed through value-added scores. For instance, in *Cook v. Stewart*, the court had to assess the state's claim that a teacher could have a causal impact on a student's learning despite never teaching the student directly but simply being in the same school building. The court considered that conclusion to be plausible, endorsing the state's causal story that "it is rational to believe 'that professional teachers could positively impact all students in their schools by fostering an encouraging learning environment, influencing and inspiring other teachers . . . or otherwise making efforts to improve overall student learning'" (*Cook v. Stewart*, 2014, p. 1213). Similarly, the plaintiffs' theory in *Vergara* relied heavily on data from teacher evaluation systems put in place by recent legislative school reform efforts. The unions opposing the case were concerned with protecting teachers from the adverse effects of these systems and hoped to convince the court of the limited utility of value-added scores for understanding the operation and quality of

public school systems. Given its focus on system-wide change, *Vergara* can also be characterized as public law litigation in key respects. Indeed, the theory of action underlying the plaintiffs' attempt to promote system-wide change in *Vergara* turns on the concept of teachers and their relationship with the education system as a whole and was inherent in their arguments and evidence they brought to bear. Given the centrality of this case to these broader trends in education policy and litigation as well as the potential for similar lawsuits in other states, the *Vergara* case proceedings warrant scholarly analysis. In particular, the fact that the case turns on an understanding of teacher work, teacher evaluation, and teacher labor markets makes the examination of their portrayal in these proceedings a crucial component of such an analysis.

Analytical Framework

Our analysis of *Vergara* focuses on how courts and the parties before them constructed narratives about the relationship between teachers, the education system in which they are situated, and student learning. Researchers interested in how governmental institutions make decisions have long focused on the construction and use of such narratives. Much of the work on causal stories has focused on the activities of the legislative and executive branches (e.g., Mills, 2007; Wilson, 2000), including in the context of education policy (e.g., Bulkley & Gottlieb, 2017; Harrison & Cohen-Vogel, 2012). Key to these stories, as Stone (1989) notes, is that they are actively constructed and contested:

Political actors . . . do not simply accept causal models . . . they compose stories that describe harms, difficulties, attribute them to actions of other individuals or organizations, and thereby claim the right to invoke government power to stop them. (p. 282)

This is no less the case for legal matters where the outcome will often turn the parties' ability to articulate and substantiate a specific kind of causal story that triggers a particular kind of legal analysis and secures a specific judicial determination. Thus, for our analysis, we build upon Stone's (1989) framework for attending to causal stories and expand it to consider how this happens in the context of the specific litigation stories presented in *Vergara* about teaching, learning, and the educational system.

In considering how to extend the framework in this way, it is important to account for how the unique formal rules of the law and organizational practices of courts shape the construction of causal stories in ways that may differ from those developed in other contexts such as a legislative session. A

great deal of research has focused on the central role of the law, itself, in influencing the ideas adopted by courts (Ewick et al., 1999). Legal scholars have increasingly underscored the particular ways in which legal rules—such as standards of evidence or definitions of harm—shape the way ideas and social practices are presented in courts (e.g., Brooks & Gewirtz, 1996; Edelman & Suchman, 1997; Gordon, 1984). Not only do legal proceedings shape the kinds of facts and stories that can be presented, but the ultimate outcome of a trial like *Vergara* is shaped by the court's own recounting of the facts and the narratives in the judicial opinion—one that gives certain facts and narratives the imprimatur of the law (e.g., Levinson, 1996; Phillips & Grattet, 2000; Wald, 1995). Paris (2010) refers to this evolution of ideas as they come into contact with the legal system as “legal translation” and describes the “conceptual and rhetorical processes through which reformers translate their values and goals into plausible legal claims and arguments. Legal translation involves, simultaneously, an appeal to legal authority and the selection and representation of ‘facts’ and ‘evidence’.” (p. 3).

One key dimension of this kind of judicial storytelling is its choice of narrative scope. Lempert and Sanders (1986), in particular, have examined how courts frame ideas in relation to the degree to which courts consider social context. They argue that when making decisions in legal cases, judges generally engage in typification—the process by which they organize and categorize experience. Through typification, two elements determine how deeply courts conduct an adjudication. First, “extensiveness” characterizes a court's willingness to consider evidence separated in space and time from the specific event in question. Second, “uniqueness” determines the “fine-grainedness” of a judicial examination or how deeply a court examines the special characteristics of situation to identify its uniqueness. Similar to “agenda setting” in legislative processes (e.g., Kingdon, 1984), these determinations in the adjudication process have important implications for what constitutes relevant evidence, what constitutes a harm, and what (if anything) constitutes a legitimate remedy in the context of the case.²

For example, when courts accept “highly typified” accounts, they generally engage in only superficial analyses of situations. Such superficial analyses can obscure the social and organizational contexts surrounding the actions of individuals. As Vaughan (1999) argues:

[T]he behavior of organizations cannot be explained solely by the actions of the individual agents but is explained more accurately as a predictable and systematic product of certain characteristics of organizational systems . . . the legal and administrative apparatus for the social control of organizations has lagged behind in acknowledging the socially organized causes of organizational misconduct. (p. 291)

Although Vaughan was particularly focused on the operation of regulatory regimes, such cautions hold true for courts as well. Indeed, the extent to which a court frames problems in education as residing within organizational, social, political, and/or economic contexts instead of at the individual level is a fundamental feature of the causal narratives that courts adopt and thus plays a crucial role in the attribution of responsibility and blame for harms. Stone (1989) describes this as the distinction between a simple causal story and a complex causal story, in which the complex story attempts to grapple with and assign blame to institutional, historical, or social contexts.

One way that litigants attempt to bring in, frame, and manage this contextual complexity is through expert testimony. This testimony provides the opportunity for litigants to introduce a wide range of views and value systems as they attempt to present facts in a light that best supports their vision of the way the world works and their theory of the case (e.g., Golan, 2004; Jasanoff, 1995). Although expert witnesses can educate judges about scientific evidence and how to interpret it (Chesler, Sanders, & Kalmuss, 1988), courts have exhibited difficulty engaging with scientific evidence (Blumenthal, 2002; Edmond, 2007; cf. Jasanoff, 1995), particularly in complex fields of public policy like education (e.g., Heise, 2002; Welner & Kupermintz, 2004).

In some cases, the introduction of quantitative evidence provides a useful means of handling this complexity. Not only does this evidence provide the appearance of neutrality—a valuable form of authority (Espeland & Stevens, 2008; Hacking, 1990; Porter, 1995)—but numbers also often provide a useful routine for dealing with complex terrain, competing value systems, or phenomenon for which value is hard to calculate. The simple presentation of numbers, however, is not enough to ensure their acceptance. Courts are often asked to determine the legitimacy, validity, and acceptable use of numbers in cases ranging from test scores (*Larry P. v. Riles*, 1979) to college admission policies (*Grutter v. Bollinger*, 2003) to actuarial practices (*Los Angeles Water and Power v. Manhart*, 1978). Indeed, although they may be contested and even outlawed, numbers are frequently important anchors for causal narratives constructed by courts. Numbers and the calculations that produce them have become important tools for evaluating—often under judicial scrutiny—public works projects (Espeland, 1998; Porter, 1995), workers' compensation insurance (Witt, 2004), and monetary value of nature (Fourcade, 2011). In the specific context of education, monetary calculations have featured prominently in the decades of school finance litigation in which judges have consistently used number of dollars and test scores as proxies for quality (Superfine, 2009; Sabel & Simon, 2004).

Similarly, the ability to quantify teacher quality—and its effects on student school and life outcomes—has become an increasingly important aspect of

education policy. Given the controversy surrounding the use of these measures, courts have been asked to weigh-in on the legitimacy and extent of their use (Superfine & Thompson, 2016; Pullin, 2013). At stake in these determinations, as we examine below, is often a court's acceptance of typified counts of teaching and learning. To the extent that courts accept quantifiable evidence as reflective of individual teacher effectiveness net of other factors, courts accept highly typified accounts that foreground the causal influence of individual teachers over any number of broader contextual factors for producing student learning.

Despite our focus on courts' construction of causal narratives, it is worth noting that courts do not work in a vacuum. Specifically in education cases, courts have looked to other governmental branches to help them work through complex and ambiguous legal matters. For example, courts ruling in school finance cases have looked to standards enacted by state legislatures to help them understand constitutional guarantees for an "adequate" education (Superfine, 2009). In several school finance cases, courts have explicitly stated that they do not have the authority or expertise to overrule legislative enactments, thus labeling determinations of a state's funding formula a "non-justiciable political question" and deferring to the judgment of the legislature (Enrich, 1995).

Courts also have demonstrated sensitivity to public opinion. On a fundamental level, courts cannot decide cases without reference to broad understandings about U.S. culture and society that are shared by a large number of people or such decisions simply would be unintelligible (Post, 2003). Welner (2012) argues that courts specifically in education cases operate in a "zone of judicial constraints" that is at least partly formed by public opinion. In other words, while courts can make decisions that go against public opinion, perception of public opinion is an important factor constraining judicial decision making. Moreover, when courts make decisions in high-profile cases, they enter into "conversations" with the public that evolve over time about the ways in which we do and should live in light of social, technological, and economic changes and in light of how such decisions are actually implemented (Kramer, 2006). While the courts may not independently undertake or sustain unpopular policies, they can still push the debate in a certain direction in the face of divided public opinion and ultimately respond to that opinion as it develops.

Such insights about courts' relationship with other governmental branches and the public highlight that the process of typification and development of causal narratives in courts have a number of potential points of contact with the development of causal stories outside of the legal process. As Stone (1989) argues, ". . . our understanding of real situations is always mediated by ideas;

those situations in turn are created, changed, and fought over in politics” (p. 282). Thus, the development of causal stories is a critical part of the policy-making process. These stories are a way for policy makers to frame contested issues, define problems, and assign blame to a range of actors in a way that is beneficial to that group of policy makers (Rochefort & Cobb, 1994; Stone, 2011). Stone (1989) describes these stories as the way certain conditions, such as poor student achievement, are redefined as problems that can be solved via policy. These stories are the ways political actors or policy makers offer their interpretation of the problem at hand, how that problem can be solved, and who is implicated in that problem. Indeed, the “conversations” about causal stories that courts have with other branches, and the public more generally, are potentially a critical factor influencing both the enactment of policy and legal decisions about important educational issues.

As we explore below, the adjudication of the *Vergara* involved the production of competing causal stories about the relationship between teachers, hiring practices, student outcomes, and the challenged statutes—one that ultimately turned on courts’ acceptance of empirical evaluations of teacher quality. Our analysis proceeds by examining the causal stories constructed by the trial and appellate court judges, how these stories were constructed from the evidence presented at trial, and the vision that emerges from these stories about the role of teachers and the operation of the California school system.

Data and Analytical Method

Our data comprise nearly a 5,000-page *Vergara* trial transcript along with the trial court opinion and the final appellate court decision (the California Supreme Court declined to hear the case). Our primary analytic method for interpreting these data is legal analysis, which involves the qualitative analysis and coding of legal proceedings. It is similar to other forms of qualitative coding but differs primarily in its explicit attention to the particular organizational, cultural, and epistemological (i.e., evidentiary) norms of legal practice (e.g., Blumenthal, 2002; Jasanoff, 1995). Recognition that legal proceedings are adversarial, outcome (i.e., verdict) driven, and structured by existing legal precedent is a crucial starting point for analyses of the trial court narratives and verdicts reached in a particular case (Kagan, 2003).

Much of the literature that examines the discourse of trial court proceedings has been conducted on jury trials (e.g., Conley, 2015; Goodwin, 1994). Often the analysis—whether in criminal or civil cases—has been aimed at examining the way jurors make sense of conflicting accounts of the same events (e.g., Bennett & Feldman, 1981; S. Harris, 2001; Hosch, Beck, & McIntyre, 1980), expert testimony at trial (e.g., Goodwin, 1994; Krauss & Sales, 2001), or

unfamiliar legal categories (e.g., Brown, 1998; Hoffmann, 1995; Severance & Loftus, 1982). *Vergara*, however, involved a facial constitutional challenge and a bench trial. We presume, therefore, that the case did not turn on the judge's familiarity with legal categories but rather on the visions of the education system underlying the competing theories of the case.³

With this in mind, we began by identifying the sections of the court transcript that involved discussions of teachers and teacher work, discussions of administrators and administrator work, and exchanges pertaining to the broader operation of the school system (e.g., district hiring practices). We began by independently coding the same portion of the trial transcripts and court opinions using an open-coding method (Glaser & Strauss, 1967; Strauss & Corbin, 1990). Drawing upon Stone's (1989) typology of causal narratives, we looked for how problems and solutions were defined and how blame and responsibility were apportioned. We used Stone's (1989) typology as a starting point, but as Stone herself stated, her typology is meant to demonstrate how political actors can develop and manipulate causal stories through their portrayal of intent, consequences, and blame. Informed by the unique aspects of causal narratives in the judicial setting, we looked for narrative scope, including how issues of teaching and teacher quality were presented and framed, what expert testimony was presented, and the extent to which quantitative evidence was presented. These categories formed our initial set of codes.

Data analysis began with identification of the sections of the trial transcript that contained arguments or expert testimony about teaching, learning, and the educational system. We used the initial set of codes to individually code the identified portions of the trial transcript. In addition to the initial set of codes, we also individually coded emerging themes and ideas. We then reconciled our emergent codes through discussion and revisiting the data until agreement across all researchers was reached.

Codes that emerged from this round of analysis included the role of the teacher in student learning, the definitions of effective or ineffective teachers and teaching, the extent to which teachers are professionals, references to teachers' unions, and references to school administrators. These collective codes were then used to code the rest of the identified sections of the trial transcripts, trial court, and appellate court opinions, in which teachers, administrators, or schools were discussed. The remaining sections were manually coded by one researcher using the set of codes developed by the researchers, along with random portions selected for dual coding to ensure reliability across researchers. Coding discrepancies were discussed collectively, and the data revisited, until consensus and the relevant portions of the transcripts were recoded. We then used the coded data to construct the narrative developed by the litigants, as well as the narratives presented in the trial and

appellate court decisions. The development and comparison of these narratives formed the basis for our conclusions in this study.

It should be noted that in limiting ourselves to an analysis of the trial court transcripts and judicial decisions, we are not suggesting that other factors beyond what was presented at trial may have entered into the judicial decision making, and our inability to address those factors is a limitation of this study. We readily acknowledge that a wide range of nonlegal factors can affect the outcome of court cases (e.g., Goldman & Jahnige, 1976; Welner, 2012), and one of the authors has explicitly considered those factors in the context of this case (Superfine & Thompson, 2016). However, we consider attempting to address how these factors may have entered the courtroom or influenced the proceedings to be beyond the scope of the present inquiry. Even without linking the narratives to external factors, the rationale produced by the courts are worthy of our attention because the rationale that produced the decision becomes constitutive of future understandings and the subject of subsequent legal and social contests over these issues. In excavating and examining the explicit rationales and visions of the school system articulated in this case, we call attention to their existence and to their potential influence on future legal disputes as well as contemporary understandings—legal and lay—of schools.

Findings

Our analysis of the trial court and appellate decisions reveals that each decision rests on a different causal narrative about the relationship between teachers, administrators, and the broader environments in which schools are situated. These narratives share core ideas about the extent to which an effective teacher can be identified and effective teaching quantified. However, these narratives differ in the extent to which the management of individual schools and school districts can ensure equitable distributions of quality teachers. In the following section, we present these causal narratives. Given that the two courts that considered the case came to different conclusions—one striking down the statutes, the other preserving them—we present the narratives in the context of the trial court and appellate decisions and examine the ways in which they draw on and reformulate the evidence (and causal narratives) presented at trial.

Trial Court Decision

The *Vergara* plaintiffs alleged that the five challenged statutes violated students' state constitutional equal protection and educational rights by allowing

“grossly ineffective” teachers to receive tenure and remain permanently employed due to elaborate dismissal procedures and teacher labor force rules that operate on a mechanical “Last-in-First-Out” rule. Constructing a narrative about the “grossly ineffective teacher”—including his prevalence and lasting effects on students—became a central focus on the trial proceedings. From the opening statements to the closing statements, both sides acknowledged that grossly ineffective teachers were currently employed in the California system (e.g., Trial Tr., 2014, pp. 12, 45-46, 4529, 4593). In describing the behaviors of the grossly ineffective teachers, both sides also used similar terms. The plaintiffs’ lawyers spoke of teachers “with no lesson plans at all” who read “a newspaper all day instead of teaching students or those who “play YouTube videos while their students sit in the classroom waiting to hear a lesson” (e.g., Trial Tr., 2014, pp. 12-13). The defense, likewise, offered the example of “those who are asleep in the classroom” as an obvious sign of a grossly ineffective teacher—making the 18 months of the tenure process “more than enough time to identify grossly ineffective teachers.” (Trial Tr., 2014, p. 45).

Both sides also agreed that identifying effective teaching was about as easy to spot as grossly ineffective teaching. During his trial testimony, then Los Angeles Unified School District (LAUSD) Superintendent John Deasy said that effective teaching could be identified via a range of components, including content knowledge, a teaching license, classroom management, and incorporating student contexts into instruction (Trial Tr., 2014, p. 110). The defense in its closing statement stressed that effective teaching could be easily identified, noting that “it’s not rocket science” and that “it doesn’t take complex mathematical equations that no one can understand to determine whether a teacher is ineffective.” Citing testimony from Linda Darling-Hammond, the defense argued that all it takes is “commitment, standards, and follow-through” involving “observations and . . . certain artifacts of the teachers work” (Trial Tr., 2014, p. 4596). The plaintiffs agreed. They specifically noted in their own closing statement suggesting that no administrators, called by either side, reported having trouble identifying effective and ineffective teachers (Trial Tr., 2014, p. 4539).

While both sides agreed on the relative ease of identifying effective and grossly ineffective teachers, they developed different narratives around the causes and solutions to the problem of ineffective teachers in the classroom. The plaintiffs hoped to use the narrative to underscore the absurdity of a set of statutes that allowed these obviously ineffective teachers to keep their jobs and, depending on their seniority, to prevent them from being fired even at a time of teacher layoffs. The defendants, by contrast, sought to emphasize that the problem was not with the statutes—the dictated order of layoffs or the

length of time to tenure, for example—but with those administrators who had missed the obvious signs of good and bad teaching in the first place.

Another crucial difference was the narratives each side crafted concerning the harms associated with the presence of “grossly ineffective teachers” and the potential gains to be had from removing them. The plaintiffs’ lawyers built their narrative around the quantitative, economic expertise of two expert witnesses Harvard economists Raj Chetty and Thomas Kane. Relying on evidence from “value-added models” that allowed them to speak hypothetically about the value of a student being taught by an average or highly effective teacher rather than an ineffective one, Chetty expressed this harm in terms of lost lifetime earnings; Kane in terms of months of school learning lost (Trial Tr., 2014, p. 469, 510, 1280). Chetty and Kane’s testimony emphasized both (and crucially for their theory of the case) that the grossly ineffective teachers were the cause of these harms to students and that the identified harm could be avoided, or at the very least mitigated, by a substitution of an average teacher for the ineffective one identified by their models. The result was the production of a causal story that framed the issue narrowly—one that began and ended with individual teachers and did not require considerations of the systematic concerns about poverty or working conditions.

The causal story presented by the plaintiffs’ experts directly sets up removing ineffective teachers as the solution to improving these outcomes. As effectiveness is a quality located within individual teachers, ineffective teachers, because they can be easily identified, should be removed from teaching and replaced with effective teachers. In this narrative, the challenged statutes, tenure, dismissal, and LIFO (last in, first out) are obvious impediments to the removal of ineffective teachers. If these statutes were removed, all students would have greater and more equal access to quality teachers by removing the constraints on districts’ abilities to remove them and by no longer requiring districts to use “LIFO” to determine which teachers to lay off.

The combination of the vision of teachers as easy to identify and the magnitude of the harm as readily quantifiable proved decisive. In ruling on behalf of the plaintiffs, the trial court judge drew directly from the evidence provided by Chetty and Kane and restated the logic of the causal narrative their evidence implied. The trial court’s decision links teacher effectiveness with a variety of measurable outcomes such as student learning or lifetime earnings: “Based on a massive study, Dr. Chetty testified that a single year in a classroom with a grossly ineffective teacher costs students \$1.4 million in lifetime earnings per classroom” (Trial Tr., 2014, p. 8). The decision also refers directly to Dr. Kane’s testimony, stating, “students in LAUSD who are taught by a teacher in the bottom 5% of competence lose 9.54 months of learning in a single year compared to students with average teachers” (p. 8). These

recapitulations of the evidence presented by Chetty and Kane in the trial court opinion make it clear that the judge accepted the causal story in which teaching, particularly truly ineffective teaching, can be identified and linked to negative effects on important student outcomes.

Absent from the narrative crafted by the trial court judge was a consideration of the broader context in which prospective teachers sought jobs and hired teachers performed their work. The defense's expert witnesses had tried to situate the potential harm of ineffective teachers in a broader frame. Jesse Rothstein, a labor economist, testified to the difficulty—conceptually and methodologically—of atomizing teachers and isolating their unique influence. In particular, he stressed the inability of the value-added models used by the plaintiffs' experts to disentangle the influence of individual teachers from other factors operating within the school such as the availability of textbooks, student mobility, administrator quality, and general school climate (Trial Tr., Transcript, 2014, pp. 2733-2735) and the inability of the models to deal with the lack of random assignment of students to schools as wealthy families frequently choose to live in areas believed to have better schools (Trial Tr., 2014, p. 2740). Inadequately dealing with nonrandom assignment would make it difficult to compare teachers' effectiveness across school settings—a claim central to the plaintiffs' theory of the case. Relatedly, Rothstein also expressed skepticism that Chetty's models demonstrated that whatever effects could be associated with the teacher could be presumed to be stable across school contexts noting that most of Chetty's examples of teacher value-added scores across contexts actually involved teachers changing grades *within the same school* (Trial Tr., 2014, pp. 2738-2739)—a disagreement that highlighted the divergent views of the relevance and characterization of even immediate context.

David Berliner, an educational psychologist, likewise testified for the defendants that school-based variables only accounted for 20% of variation in student learning, while those out of school factors—low birth weight, lack of medical care, food insecurity, family stress (Trial Tr., 2014, pp. 3800-3803)—count for far more. Even considering the in-school effects, Berliner noted that teachers only account for roughly half of the in-school variation—underscoring his general view “that teachers are accounting for far less of the variation we see in scores than is generally believed, and out-of-school factors are accounting for far more of the variance in scores than people generally believe” (Trial Tr., 2014, pp. 3804, 3811).

Although this narrative in which teacher effects are situated in the larger context of the school and the communities in which students live and teachers work was offered at trial, there is no evidence of it in the ruling by the trial court. There is no mention of the effects of nonrandom assignment of

students to teachers, the effects of nonteacher related in school factors like school climate, or the effects of nonschool factors like poverty. Instead, the trial court judge determined that the challenged statutes necessarily deprived students of their constitutional rights to a quality of education because the statutes “result in grossly ineffective teachers obtaining and retaining permanent employment” and because of their disproportionate assignment to schools with large low-income and minority student. The court also implicitly embraced the hypothetical implied by the analysis of the plaintiffs’ experts: the ready availability of a pool of untapped teachers of average quality waiting to take the place of those deemed grossly ineffective—thereby addressing the constitutional deprivation. With these determinations of fact, the judge gave legal weight to a specific causal narrative of student achievement rooted in a specific view of teachers, teaching, the teacher labor market, and their effects on students’ life trajectories. As with all causal narrative, this one implied a specific solution and form of intervention. If large numbers of easily identifiable, grossly ineffective teachers damage the educational and life opportunities of students, then they can and should be removed by administrators. Any laws that prevent administrators from doing so are inherently suspect and, in the trial court judge’s view, also unconstitutional.

In short, the causal narrative presented in the trial court decision is a simple, clear story, in which individual teachers are identifiable as effective or ineffective teachers, their effects on student outcomes can be quantified, and removal of ineffective teachers will lead to improved student outcomes. Missing from this narrative is any discussion or substantive treatment of elements outside of the teacher’s control, such as student life circumstances, school and district management, and school funding not positioned as factors that should be taken into consideration when making personnel decisions. Although witnesses on both sides acknowledged to varying degrees the importance of external school factors, they do not feature in the trial court opinion—appearing only obliquely in the statement “all sides . . . agree that competent teachers are a critical, if not the most important, component of success of a child’s in-school educational experience” (Trial Tr., 2014, p. 7). How these implied *out of school* factors might constrain, augment, or complicate the causal narrative of the opinion is left unstated.

Appellate Court Decision

The causal story presented in the appellate decision shares with the trial court decision the position that teachers are an important factor in student learning, and that highly ineffective teachers will have a negative effect on student

outcomes. Both narratives also share the idea that effective and ineffective teachers can be identified. Although it reached a different legal conclusion, the appellate court decision did not challenge the idea that teachers are essentially effective or ineffective nor did it complicate the idea of teacher effectiveness by, for example, exploring the ways in which effective teaching may be context-specific. Although the narrative in the appellate court decision does not emphasize the quantifiable nature of the effects of teaching in the way the narrative in the trial court decision does, it shares the basic premises that teacher quality is knowable, that teachers are central to student learning, and that poor teaching will hurt the life opportunities of students remained unchallenged (*Vergara v California*, 2016, p. 35).

While the two opinions share these basic elements of their causal narratives, the appellate decision positioned these elements in a different story about the relationship between the challenged statutes and improved teacher quality. Where the trial court decision posited that the challenged statutes were the direct impediment to the removal and replacement of ineffective teachers that would lead to more equal educational opportunities, the appellate decision concluded that the impediment existed not at the level of state statutes but at the level of *school administration*.

The appellate court's causal story framed the problem of ineffective teachers as a failure of district officials to perceive teacher quality correctly, to make correct hiring/retention decisions, and to distribute teachers equitably. For example, the appellate decision states:

... the evidence at trial showed what the text of the challenged statutes makes clear—that the challenged statutes do not in any way instruct administrators regarding which teachers to assign to which schools. Thus, it is administrative decisions (in conjunction with other factors), and not the challenged statutes, that determine where teachers are assigned throughout a district. (*Vergara v California*, 2016, p. 33)

The appellate court notes that this does not mean that poor teachers are not inequitably distributed to low-income or minority schools but asserts a different cause for their presence in those schools: not the statutes and hand-cuffed administrators of the trial court's narrative but rather the poor judgment and personnel practices of some school administrators.

The appellate decision repeatedly tells a story that assigns dispositive agency to district administrators rather than to the statutes and, in turn, the individual “grossly ineffective teachers” the trial court argued the statutes protect. Noting, for example, “administrators—not the statutes—ultimately determine where teachers within a district are assigned to teach”

and, specifically critique the trial court's specific causal claim and proposed remedy, "even if the statutes were struck down, the harm at issue . . . could still occur as before" (*Vergara v California*, 2016, p. 6, 35).

The appellate court went further to illustrate the potential effects of administrators' agency noting the negative consequences of the "deplorable staffing decision made by some local administrators that have a deleterious impact on poor and minority students in California's public schools" (*Vergara v California*, 2016, p. 35). Supporting this observation, the court specifically cited the positive consequences noted in the testimony of former LAUSD Superintendent John Deasy's testimony concerning the district's move away from a mechanical, "virtually automatic" tenure system to an active, agentive system involving a thorough review of a teacher's abilities (*Vergara v California*, 2016, p. 18). The court additionally noted the "various strategies for resolving dismissals short of the formal dismissal process" (*Vergara v California*, 2016, p. 19) used by school administrators called by the defense—an acknowledgment that an administrator's agency arsenal is multifaceted involving both hard and soft power. The court buttressed these observations by citing the testimony of defense expert witness Linda Darling-Hammond who observed that the short time to tenure in the Challenged Statutes actually encouraged the proactive use of administrator agency, and any change might have the unintended consequence of encouraging passivity and delaying consequential actions (*Vergara v California*, 2016, p. 18).

Implied in this narrative is that the failure to have effective teachers in all classrooms is not the result of an inability to remove teachers due to arduous bureaucratic processes specified in the challenged statutes, but a failure of district officials to make appropriate use of the discretion afforded to them in staffing decisions. In this story, the cause of the identified inequity is assigned to district officials, rather than individual "grossly ineffective." As in the trial court decision, though, limited consideration is given to factors that influence the immediate hiring decisions of school officials. Like the trial court, the appellate court's decision was premised on a view that there were additional effective teachers available in the labor market. While the trial court was concerned by the inability of administrators to hire those available effective teachers because of the Challenged Statutes, the appellate court essentially faulted administrators for not continuing to seek them out but instead settling for those ineffective teachers once chosen. The appellate court did not concern itself with the possible constraints administrators might face in securing and maintaining a steady supply of effective teachers, the cost-benefit trade-offs associated with increased teacher turnover, or with the possibility that a previously effective teacher might cease to be effective if placed in a different school or community context.

The causal narrative told by the appellate court's decision shares a similar structure to the trial court's causal narrative. In both cases, the ability to identify effective and ineffective teacher and to measure their impact on students is unchallenged. Both narratives also share a concern for the effect of the disproportional assignment of ineffective teachers to low-income and minority students. While these narratives differ in the solutions they offer to this problem, both narratives share a simple structure, in which all students can have access to equal educational opportunities, provided they are assigned an effective teacher and in which larger social conditions and external factors do not appreciably constrain the ability of administrators to find good teachers and secure more equitable educational outcomes.

The Road Not Taken

One striking similarity between the trial and appellate court causal narratives is the extent to which these narratives fail to reference the social, political, cultural, or economic context of teaching and learning. The two narratives provided by the courts present a story in which unequal educational opportunities can be addressed through the correct assignment of quality teachers, through either improved statutes or better administrative decision making. Thus, these narratives share a core assumption, evident in their simple structure, that unequal educational opportunities stem from ineffective teaching, rather than the web of social, political, and economic opportunities that varies across contexts.

The appellate court decision gestures at a more complex story by touching upon the impact of teacher quality relative to other factors on student achievement. For example, the appellate court acknowledges the role of student context in achievement, stating, "although a host of factors, including child poverty and safety, affect student achievement, teachers nevertheless have a highly important and significant impact on student learning" (*Vergara v. California*, 2016, p. 13). The court also highlights testimony from David Berliner, in which he argued that teacher quality accounted for approximately 10% of variation in student test scores, and thus the majority of variation in test scores can be assigned to factors other than teacher quality (*Vergara v. California*, 2016, p. 17). Both of these details imply that the court acknowledged that schools were embedded in very complex environments and engaged in an endeavor of multiple and uncertain causality.

The appellate court also took judicial note of several "extremely troubling" dynamics stemming from teacher labor markets and personnel practices but stopped short of connecting this acknowledged complexity to their basic theory of how low-income and minority students are disproportionately assigned to ineffective teachers. The court acknowledged the disturbing

nature of this phenomenon, describing the dance of the lemons as a practice that “should not be allowed to occur” but that it is ultimately a process “driven by local administrators” (*Vergara v. California*, 2016, p. 34). The court notes specifically that these administrators “rid their schools of highly ineffective teachers by transferring them to other schools, often low-income schools.” The court also references testimony from expert witnesses for the plaintiffs, which stated that grossly ineffective teachers “tend to” cluster in schools that serve low-income and minority students. Noting that some collective bargaining agreements allow teachers to select where they want to teach, the court stated, “Teachers with such an option often chose not to teach at schools predominantly serving low-income and minority children because the schools can have challenging working conditions” (p. 17). Similarly, the appellate court acknowledged that when teachers can choose their assignments, teachers often choose not to work in schools serving predominantly low-income and minority students due to the “challenging working conditions” of those schools (*Vergara v. California*, 2016, p. 17).

The deeper political, economic, and organizational questions of *why* these administrators might choose to transfer these teachers as opposed to firing them or why low-income schools might accept (willingly or not) the transfer of these teachers or why teachers might voluntarily and en masse decide not to teach in schools populated predominately by low-income and minority students is left largely to the imagination. Whatever the reasons for these administrators’ and teachers’ decisions and whatever the reasons for the court’s failure to inquire after them, the effect is to preserve administrators’ agency in the causal story: Administrators should not engage in the dance of lemons by transferring or accepting the transfer of effective teachers and, presumably, they should not negotiate collective bargaining agreements that allow teachers to select out of low-performing schools.

Although neither the trial court nor the appellate court engaged in a discussion of the broader contexts of schools or the lives of the children who attend them, it was not because the evidence was not included in the trial court record. Indeed, during the trial at least two expert witnesses explicitly discussed the relationship between student context, particularly living in poverty, and the working conditions in a school. Susan Moore Johnson testified that while teachers, on average, reported better working conditions in affluent schools, the working conditions, particularly social conditions such as school culture, account for a great deal of the perceived relationship between student demographics and teacher turnover (*Vergara v. California*, 2014, p. 2018). She went on to argue that the challenged statutes help create positive working conditions for teachers through tenure and a formalized dismissal process. Kevin Futernick, another witness for the defense, argued one central

challenge facing schools with large populations of low-income students was the adverse effects of living in poverty and the ways in which these effects could harm the working conditions of the school, and thus encourage teachers to leave the school (*Vergara v. California*, 2014, p. 4006).

Despite acknowledging that grossly ineffective teachers are disproportionately assigned to schools that serve low-income and minority children, that these schools often have challenging working conditions, and the availability of expert testimony to this effect, the courts did not take the additional analytical step to explore why this was the case. The trial court obviated the need for this inquiry by settling on a theory of teacher quality identification and labor market supply that made this analysis extraneous; the appellate court simply found that the plaintiffs failed to prove that this was the inevitable result of the challenged statutes and thus failed to demonstrate a facial constitutional violation. In doing so, both courts opted for simpler causal narratives, by ignoring the complicated social factors that gave rise to those challenging working conditions. In this respect, both courts developed a highly typified account of the relationship between teaching and learning, offering a causal story in which equal educational opportunities are a function of getting effective teachers in all classrooms, rather than a function of social context writ large.

Discussion and Conclusion

Shared Views, Divergent Narratives

Our analysis of the trial and appellate court decisions indicates that both judges engaged in causal storytelling and, in the process, provided highly typified accounts of the relationship between student learning, teacher practice, and the general operation of California's school system. The plaintiffs' legal theory relied on the court accepting a tight causal chain that linked, first, ineffective teaching to poor student school (and life) outcomes and, second, linked the prevalence of low-quality teachers to the limited degrees of freedom available to administrators in the hiring and firing of teachers. According to the plaintiffs, the mechanistic decision-making process demanded by the challenged statutes—the receipt of tenure after 2 years, the LIFO dismissal rules—inevitably led to the deprivation of Californian students' constitutional rights. This is precisely the view adopted in the trial court decision. Striking down the laws would remove the “unguided,” to use Stone's (1989) term, portion of the teacher selection system, in which teacher placement is not the direct result of any intentional actions, and restore the agency of principals and administrators to guide the selection of effective teachers and, in turn, improve student outcomes.

The appellate court's ultimate legal conclusion did not *require* the court to engage in causal storytelling. While the plaintiffs needed to demonstrate that the challenged statutes inevitably resulted in the described violations of the equal protection rights of low-income and minority students through the construction of a causal story that linked the violations to actions that stemmed directly from the challenged statutes, the appellate court simply had to find reasons *other than the challenged statutes* for the violations—a task that did not require developing specific causal stories rooted in the actions of specific school officials or teachers. Nevertheless, the court's opinion nevertheless engaged in causal story telling—one that mirrored the trial court's story in important respects. Specifically, the appellate court accepted the first link in the causal chain connecting teacher effectiveness to student outcomes. The appellate court, however, told a different story when it came to the relationship between the Challenged Statutes and the prevalence of ineffective teachers in certain schools. Rejecting the trial court's characterization of the statutes as producing a mechanical, unguided hiring process, the appellate court argued that administrators retained considerable discretion and control. According to the appellate decision, it was the differentially effective use of this discretion and not the Challenged Statutes that were responsible for the inequitable outcomes identified by the plaintiffs' experts.

While the two courts told different causal stories and disagreed on the extent to which the challenged statutes produced mechanistic hiring and retention practices, both courts produced highly typified accounts of California schools. Consistent with the law and society literature (Lempert & Sanders, 1986; Vaughan, 1999), the judges minimized the extent to which they considered the broader social context in which teachers, administrators, or schools operate. This is most evident in the centrality of value-added scores to the reasoning of both opinions. Value-added scores, in purporting to provide an individual teacher's contribution to student learning independent of school and social context, provide a paradigmatic example of typification. Consistent with prior research on the value of quantification in general and quantified representations of schooling in particular (e.g., Dunn & West, 2008; Espeland & Stevens, 2008; Posselt, 2016; Rivera, 2011), the value-added scores provided a basic foundation for the construction of causal stories. In the case of the trial court, the inability of principals to maximize the average value-added scores of teachers or to utilize the most reliable measure of a teacher's effectiveness was key to substantiating the constitutional violation.

The appellate court did not challenge the basic utility of value-added scores, nor did it try to complicate the notion that value-added scores provided a universal—that is, context neutral—assessment of teacher quality.

Indeed, it was central to the court's theory that inequity stemmed from deliberate administrative decisions and not the statutes: "Assuming that poor and minority students encounter more grossly ineffective teachers and that this impacts their constitutional right to 'basic educational equality', the constitutional infringement is the product of staffing decisions, not the challenged statutes" (*Vergara v. California*, 2016, p. 35). This view only makes sense if one accepts the premise that teacher quality affects student outcomes and is knowable independent of context. That is, a judge could hardly chastise administrators for inequitable assignment unless one believed that a teacher would retain his effectiveness if that teacher was given different students, in a different school, in a different social context.

The highly typified accounts, or simple causal stories, crafted by both courts provide a circumscribed view of the school and societal contexts in which student learning is situated. The result was that the judges did not consider the ways in which their causal stories might be constrained—and their imagined remedies inhibited—by these broader circumstances. For instance, in suggesting that student outcomes could be improved by allowing for easier removal of ineffective teachers, the trial court judge did not consider the available supply of more effective teachers or the extent to which this supply is shaped by external factors like geography, school climate, or individual teacher preferences. Indeed, the empirical literature suggests that there is complex relationship between the ability to hire, prepare, and retain high-quality teachers and the teacher supply (Bracey & Molnar, 2003).

Similarly, the appellate court implied that many of the harms alleged by the plaintiffs could be mitigated through more equity-minded decision making by administrators and more widespread use of informal dismissal strategies like counseling out (*Vergara v. California*, 2016, p. 19). While judicial recognition of these strategies provided a sufficient basis to deny the inevitability of a constitutional violation, the appellate court did not contemplate the very real limits of these strategies: Counseling out is only effective if there is an available supply of more effective teachers willing to teach in a particular school; transferring ineffective teachers to other schools provides a solution at the level of individual schools but not necessarily at the level of districts or the state as a whole. Similarly, the court chose not to delve into available testimony regarding the impact of poverty on student learning, teaching, and working conditions. The appellate decision does appear to glimpse this limitation noting that "some principals rid their schools of highly ineffective teachers by transferring them to other schools" as evidence of the agency available to principals while at the same time characterizing the practice as "extremely troubling" (*Vergara v. California*, 2016, p. 34). The court did not, however, extend its analysis beyond this negative characterization.

Agency and Inevitability

As many legal scholars have noted, the stories constructed by courts are a by-product of the relationship between legal procedure, judicial precedent, and the specific facts available and presented in the case. This makes them prone to vary from the kinds of stories that social scientists or the lay public might tell about, say, the relationship between teacher hiring practices, student achievement, and educational equity. In our analysis, the *Vergara* case appears to be no exception. While the appellate court and the trial court differed in their assessment of how much agency is retained by school administrators under the Challenged Statutes, neither constructed a story that embraces the kind of complexity or multicausality that social scientists who research these subjects are accustomed to telling, such as the complex causal stories described by Stone (1989). Although these issues were raised on both direct and cross-examination by experts witnesses during the course of the trial to acknowledge, for example, that teachers were the most important *in-school* factor—not overall factor—affecting student achievement (e.g., Trial Tr., 2014, p. 3804); that policy environment and school context matter (e.g., Trial Tr., 2014, p. 2734); and that teacher labor market preferences matter (e.g., Trial Tr., 2014, p. 2706), these factors did not substantively figure into either judicial narrative. The result, in our view, are narratives that take an all-or-nothing view to system agency rather than consider the ways in which actors at all levels—teachers, principals, superintendents—are enabled and constrained by system factors.

The appellate court, for its part, did appear to acknowledge that it was hamstrung in its reasoning by the judicial precedent imposed by the plaintiffs' choice of legal challenge—pointing to the limits of courts as conduits of reform. In several places, the court's decision seems to imply that the plaintiffs erred in their theory of the case. The court noted, "Plaintiffs elected not to target local administrative decisions and instead opted to challenge the statutes themselves" (*Vergara v. California*, 2016, p. 35)—a decision that required the court to determine that the violation stemmed "inevitability" from the statutes. The court also opined that even under their chosen theory, the plaintiffs could have prevailed if they had emphasized how *implementation* decisions—rather than the statute text—inevitably caused minority students to receive an inadequate education. The court observed that

it is possible, though not certain, that plaintiffs could have made such a showing by providing that any implementation of the statutes inevitably resulted in the consequential assignment of disproportionately high numbers of grossly

ineffective teachers to schools predominantly serving low-income and minority students. (*Vergara v. California*, 2016, p. 32)

Given the court characterized the evidence presented at trial as having “highlighted likely drawbacks to the current tenure, dismissal, and layoff statutes” and as having “revealed deplorable staffing decisions being made by some local administrators that have a deleterious impact on poor and minority students” (*Vergara v. California*, 2016, p. 35)—decisions that were “extremely troubling and should not be allowed to occur” (*Vergara v. California*, 2016, p. 34)—it seems plausible that the court included these observations to express its frustration at being hemmed in by the plaintiffs’ legal theory of the case.

This frustration, though, only serves to underscore the limits of judicial frameworks for dealing with the complexity and indeterminacy inherent to school reform efforts. The appellate court noted that, though the *Vergara* theory did not meet this standard, it was not an impossible standard to reach in the context of education cases. The court pointed to several instances of education cases that met the inevitability standard—“when the constitutional violation flows ‘inevitably from the statute, not the actions of the people implementing it’.” But the examples provided by the court to illustrate instances where statutory schemes presented insurmountable obstacles to school official and teacher capacity to fulfill their constitutional obligations to their students are highly dubious from a social science perspective. For instance, the court, citing *Serrano I* and *Sanchez*, provides the example of a state funding statute that inevitably violates equal protection. The court, citing *Butt*, also provides the example of disparities in school term length as resulting in an obvious constitutional violation. While the court can assert that these disparities are, by definition, constitutional violations, the court would be hard-pressed to substantiate empirically the claim that the funding or term length disparities *inevitably* result in differential outcomes. Indeed, the research on these matters—especially the effects of school funding (e.g., Cohen, Raudenbush, & Ball, 2003)—is highly contested.

More than anything, these examples would seem to suggest how much courts must rely on typified accounts of school operation to reach conclusions that provide the relief (and reform) sought by plaintiffs. This is no less true if we consider the alternative standard demonstrating that a disparity occurs regardless of how the statute is implemented. Although the appellate court acknowledged in a footnote that it does not have the responsibility of determining how the plaintiffs could meet this burden, it is difficult to imagine how they could do so without relying heavily on a quantification device, like value-added scores, that provides a framework for universal, generalizable comparisons and

specific causal claims. Moreover, assertions about inevitability would require a denial of agency to many actors in the school system—superintendents, principals, parents, students—that cuts against common sense, research, and the philosophy of accountability that is the hallmark of the past three decades of school reform efforts. In addition, as Harrison and Cohen-Vogel (2012) argue, a more complex, which is to say a less typified, causal story expands the group of actors implicated, which can disrupt narratives underlying various reform strategies (e.g., test-driven accountability) that vest responsibility for a problem largely with one group and create substantial obstacles for the courts to provide the requested reforms.

Implications

As our examination of *Vergara* is aimed at illuminating how courts and parties construct and use causal narratives, this examination holds important implications for these actors. On one hand, courts might simply avoid the difficulties raised by accepting complex stories in Equal Protection cases like *Vergara* by choosing to use the rational basis test. When a court uses this test, as the appellate court did, it is very difficult to rule that a governmental action is unconstitutional. As a result, the court would be relieved of crafting a remedy in a situation that is difficult for it to understand and address through traditional legal frameworks. Similarly, a court could rule that problems that involve such complex stories are “non-justiciable political questions,” indicating that a problem is too complex for a court to address and more appropriate for a legislature. Several courts in school finance reform cases deemed the issues before them to fall into this category and accordingly left the issue for state legislatures to address (Enrich, 1995). Such responses would at least reflect sensitivity to the difficulties raised by the limited institutional capacities of courts to engage effectively with complex educational problems.

On the other hand, plaintiffs seeking reform could urge courts to accept complex stories precisely because they believe that complex problems demand complex solutions. It is axiomatic that courts are to craft remedies that match specific rights violations presented to them (Mayrowetz & Lapham, 2008). In Equal Protection cases like *Vergara*, hewing such a path could be very difficult. First, a court would need to decide that the Strict Scrutiny test applies, as the trial court did. As discussed above, strict scrutiny requires governmental action to be “narrowly tailored” to a “compelling governmental interest.” It would be very difficult, however, for any proposed governmental action aimed at remedying a problem to meet that standard—there is potentially an inherent tension between the notion of a highly specific remedy that is narrowly tailored and the type of flexible remedy needed to

address a complex problem like unequal teaching and learning in a large urban school district. That said, the concept of narrowly tailored is fundamentally grounded in the idea of fit between governmental action and goal. If plaintiffs can convince the court that there is an equal protection violation for a complex problem like teaching and learning, and that this problem is grounded in the failure of government to embrace complexity in its policies, the courts could be deployed as a useful institution for promoting reform in teaching and learning.

Clearly, there is no easily replicable model for a court engaging in such activity. The history of education reform focused on promoting more equal educational opportunities is replete with examples of both judicial and legislative failure (Superfine, 2013). There are, of course, various promising examples of judicial and legislative success, but they are far from the norm. For example, some courts have acted as “destabilization” agents that force initial change in a system with a ruling and then actively and continually oversee other more competent entities (such as legislatures and districts) as they craft solutions to the violation (Sabel & Simon, 2004). Indeed, researchers and policy makers should deeply examine such cases for a better understanding of the concrete institutional strategies and situations in which the courts can serve as productive agents for such reform. In the specific case of the courts, accepting complex causal stories and acting on them is a particularly acute problem because of processes such as typification that are fundamental to the judicial institution. However, recognizing that the most important problems of education, such as that involving teaching and learning, are complex is likely a critical first step for courts to promote fundamental reform in such areas.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. There is an on-going debate about how faithfully districts are implementing state laws requiring that teacher evaluation be linked to student test scores. For instance, Students Matter, the organization that helped file the *Vergara* case, also filed a lawsuit alleging that 13 districts are in violation of California’s Stull Act by not meaningfully including test scores in their teacher evaluations. The lower court, however, dismissed the complaint (*Doe v. Antioch*, 2016).

2. We do not mean to imply here that the judicial decision-making process is linear—making a conscious decision to typify and then—based on the perceived extensiveness and uniqueness—engage the appropriate evidence and ultimately reach a decision on the facts of the case. It is equally possible the judge begins with a conclusion in mind and works “backward” through the typification process. In either case, the determination made around these issues shape the ultimate outcome of the case and, in turn, which narratives enshrined in the law.
3. Facial constitutional challenges involve a consideration of the text of laws rather than a judgment about their implementation in practice. This judicial test—that the law is unconstitutional on its face—makes the judge’s understanding of “how the world works” decisive in determining the outcome because it requires the judge to divine a universal principle rather than an interpretation of how the law played out in a specific instance. That is, the violation must be considered an “inevitable” outcome of the implementation of the law.

ORCID iD

Ethan L. Hutt  <https://orcid.org/0000-0002-4044-3506>

References

- American Educational Research Association. (2015). AERA statement on use of value-added models (VAM) for the evaluation of and educator preparation programs. *Educational Researcher*, 44, 448-452.
- Baker, A. (2014, July 3). Lawsuit challenges New York’s teacher tenure laws. *The New York Times*. Retrieved from <http://www.nytimes.com/2014/07/04/nyregion/lawsuit-contests-new-yorks-teacher-tenure-laws.html>
- Baker, B. D., Oluwole, J., & Green, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the Race to the Top era. *Education Policy Analysis Archives*, 21(5), 1-71.
- Bennett, W. I., & Feldman, M. S. (1981). *Reconstructing reality in the courtroom*. New Brunswick, NJ: Rutgers University Press.
- Blumenthal, J. A. (2002). Law and social science in the twenty-first century. *Southern California Interdisciplinary Law Journal*, 12, 1-53.
- Boutrous, T. J. (2014, January 28). California kids go to court to demand a good education. *The Wall Street Journal*. Retrieved from <http://www.wsj.com/articles/SB10001424052702303553204579347014002418436>
- Bracey, G. W., & Molnar, A. (2003). *Recruiting, preparing and retaining high quality teachers: An empirical synthesis*. Tempe: Education Policy Studies Laboratory, Arizona State University.
- Brooks, P., & Gewirtz, P. D. (1996). *Law’s stories: Narrative and rhetoric in the law*. New Haven, CT: Yale University Press.
- Brown, D. K. (1998). Plain meaning, practical reason, and culpability: Toward a theory of jury interpretation of criminal statutes. *Michigan Law Review*, 96, 1199-1268.

- Bulkley, K.E., & Gottlieb, J.J. (2017). Policy images of teachers: How influential actors construct images of teachers. *Teachers College Record*, 119(4), 1-34.
- Chayes, A. (1976). The role of the judge in public law litigation. *Harvard Law Review*, 89, 1281-1316.
- Chesler, M. A., Sanders, J., & Kalmuss, D. S. (1988). *Social science in court: Mobilizing experts in the school desegregation cases*. Madison: University of Wisconsin Press.
- Cochran-Smith, M. (2001). The outcomes question in teacher education. *Teaching and Teacher Education*, 175, 527-546.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25, 119-142.
- Cole, D. (2016). *Engines of liberty: The power of citizen activists to make constitutional law*. New York, NY: Basic Books.
- Conley, R. (2015). *Confronting the death penalty: How language influences jurors in capital cases*. New York, NY: Oxford University Press.
- Cook v. Stewart, 2014 WL 2959248 (N. D. Fl. Gainesville, May 6, 2014).
- Corcoran, S. P. (2007). Long-run trends in the quality of teachers: Evidence and implications for policy. *Education*, 2, 395-407.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1). Retrieved from <https://epaa.asu.edu/ojs/article/view/392>
- Doe v. Antioch, Case No. MSN15-1127 (S.C. of California, Contra Costa County, 2016)
- Dunn, J., & West, M. (2008). Calculated justice: Education research and the courts. In F. Hess & L. McDonnell (Eds.), *When research matters: How scholarship influences education policy* (pp. 155-176). Cambridge, MA: Harvard Education Press.
- Edelman, L. B., & Suchman, M. C. (1997). The legal environments of organizations. *Annual Review of Sociology*, 23, 479-515.
- Edmond, G. (2007). Supersizing Daubert: Science for litigation and its implications for legal practice and scientific research. *Villanova Law Review*, 52, 857-924.
- Edwards, H. S. (2014, October 30). The war on teacher tenure. *Time*. Retrieved from <http://time.com/3533556/the-war-on-teacher-tenure/>
- Enrich, P. (1995). Leaving equality behind: New directions in school finance reform. *Vanderbilt Law Review*, 48, 101-194.
- Espeland, W. N. (1998). *The struggle for water: Politics, rationality, and identity in the American Southwest*. Chicago, IL: The University of Chicago Press.
- Espeland, W. N., & Stevens, M. L. (2008). A sociology of quantification. *European Journal of Sociology*, 49, 401-436.
- Ewick, P., Kagan, R. A., & Sarat, A. (Eds.). (1999). *Social science, social policy & the law*. New York, NY: Russell Sage Foundation.
- Forslund v. Minnesota, 62-CV-16-2161 (MN Dist. Ct. 2016).
- Fourcade, M. (2011). Cents and sensibility: Economic valuation and the nature of "nature." *American Journal of Sociology*, 116, 1721-1777.

- Giersch, J. (2014). Aiming for giants charter school legislation and the power of teacher unions. *Education and Urban Society*, 46, 653-671.
- Glaser, B., & Strauss, A. (1967). *The discovering of grounded theory*. Hawthorne, NY: Aldine.
- Glazer, S., Goldhaber, D., Loeb, S., Raudenbush, S., Staiger, D. O., Whitehurst, G. J., & Croft, M. (2011). *Passing muster: Evaluating teacher evaluation systems*. Washington, DC: Brookings Institution.
- Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Golan, T. (2004). *Laws of men and laws of nature: The history of scientific expert testimony in England and America*. Cambridge, MA: Harvard University Press.
- Goldman, S., & Jahnige, T. P. (1976). *The federal courts as a political system* (2nd ed.). New York, NY: Harper & Row.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96, 606-633.
- Gordon, R. W. (1984). Critical legal histories. *Stanford Law Review*, 36, 57-125.
- Green, P. C., Baker, B. D., & Oluwole, J. (2012). The legal and policy implications of value-added teacher evaluation policies. *Brigham Young University Education and Law Journal*, 2012, Article 2.
- Grutter v. Bollinger 539 U.S. 306 (2003).
- Hacking, I. (1990). *The taming of chance*. Cambridge, MA: Cambridge University Press.
- Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annual Review of Economics*, 4, 131-157.
- Harris, D. (2011). *Value-added measures in education: What every educator needs to know*. Cambridge, MA: Harvard Education Press.
- Harris, S. (2001). Fragmented narratives and multiple tellers: Witness and defendant accounts in trials. *Discourse Studies*, 3, 53-74.
- Harris v. Quinn, 573 U.S. (2014).
- Harrison, C., & Cohen-Vogel, L. (2012). The politics of teacher reform in Florida: Analyzing causal narratives surrounding state adoption of performance-based evaluations, performance pay, and tenure elimination. *Peabody Journal of Education*, 87, 517-534.
- Heise, M. (2002). The past, present, and future of empirical legal scholarship: Judicial decision making and the new empiricism. *University of Illinois Law Review*, 2002, 819-850.
- Hoffmann, J. L. (1995). Where's the buck? Juror misperception of sentencing responsibility in death penalty cases. *Indiana Law Journal*, 70, 1137-1160.
- Hosch, H. M., Beck, E., & McIntyre, P. (1980). Influence of expert testimony regarding eyewitness accuracy on jury decisions. *Law and Human Behavior*, 4, 287-296.
- Hutt, E. L. & Tang, A. (2013). The new education malpractice litigation. *Virginia Law Review*, 99, 420-491.
- Janus v. American Federation of State, County, and Municipal Employees, Council 31, 585 U.S. (2018)
- Jasanoff, S. (1995). *Science at the bar*. Cambridge, MA: Harvard University Press.

- Kagan, R. A. (2003). *Adversarial legalism: The American way of law*. Cambridge, MA: Harvard University Press.
- Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27, 615-631.
- Kennedy, M. (2010). *Teacher assessment and the quest for teacher quality: A handbook*. San Francisco, CA: Jossey-Bass.
- Kingdon, J. W. (1984). *Agendas, alternatives, and public policies*. Boston, MA: Little, Brown.
- Komesar, N. (1994). *Imperfect alternatives: Choosing institutions in law, economics, and public policy*. Chicago, IL: The University of Chicago Press.
- Koski, W. S. (2012). Teacher collective bargaining, teacher quality, and the teacher quality gap: Toward a policy analytic framework. *Harvard Law & Policy Review*, 6, 67-90.
- Kramer, L. D. (2006). Popular constitutionalism, circa 2004. *California Law Review*, 92, 959-1011.
- Krauss, D. A., & Sales, B. D. (2001). The effects of clinical and scientific expert testimony on juror decision making. *Psychology, Public Policy, and Law*, 7, 267-310.
- Larry P. v. Riles, 495 F. Supp. 926 (N.D. Cal. 1979).
- Layton, L. (2014, January 26). Teacher tenure goes on trial in Los Angeles courtroom. *The Washington Post*. Retrieved from https://www.washingtonpost.com/local/education/teacher-tenure-goes-on-trial-in-los-angeles-courtroom/2014/01/26/b6574ed0-8545-11e3-8099-9181471f7aaf_story.html
- Lempert, R., & Sanders, J. (1986). *An introduction to law and social science*. Philadelphia: University of Pennsylvania Press.
- Levinson, S. (1996). The rhetoric of the judicial opinion. In P. Brooks & P. D. Gewirtz (Eds.) *Law's stories: Narrative and rhetoric in the law*. New Haven, CT: Yale University Press (pp. 187-205).
- Los Angeles Dept. of Water and Power v. Manhart, 435 U.S. 702 (1978).
- Malin, M. H. (2012). Sifting through the wreckage of the tsunami that hit public sector collective bargaining. *Employee Rights and Employment Policy Journal*, 16, 533-557.
- Martinez v. State of New Mexico, complaint (N.M. Dist. Ct. filed April 1, 2014).
- Mayrowetz, D., & Lapham, J. F. (2008). "But we're in a court of law. We're not in a legislature." The promise and pitfalls of educational policy reform through the judicial branch. *Educational Policy*, 22, 379-421.
- McGuinn, P. (2010). *Creating cover and constructing capacity: Assessing the origins, evolution, and impact of Race to the Top*. Washington, DC: American Enterprise Institute.
- McGuinn, P. (2012). Stimulating reform: Race to the Top, competitive grants, and the Obama education agenda. *Educational Policy*, 26, 136-159.
- Mills, M. R. (2007). Stories of politics and policy: Florida's higher education governance reorganization. *The Journal of Higher Education*, 78, 162-187.

- National Council on Teacher Quality. (2013). *State of the states 2013: Connect the dots: Using evaluations of teacher effectiveness to inform policy and practice*. Washington, DC: Author.
- National Research Council. (2010). *Getting value out of value-added: Report of a workshop*. Washington, DC: The National Academies Press.
- Paris, M. (2010). *Framing equal opportunity: Law and the politics of school finance reform*. Palo Alto, CA: Stanford University Press.
- Phillips, S., & Grattet, R. (2000). Judicial rhetoric, meaning-making, and the institutionalization of hate crime law. *Law & Society Review*, 34, 567-606.
- Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton, NJ: Princeton University Press.
- Posselt, J. R. (2016). *Inside graduate admissions: Merit, diversity, and faculty gate-keeping*. Cambridge, MA: Harvard University Press.
- Post, R. C. (2003). Fashioning the legal constitution: Culture, courts, and law. *Harvard Law Review*, 117, 4-112.
- Pullin, D. (2013). Legal issues in the use of student test scores and value-added models to determine educational quality. *Education Policy Analysis Archives*, 21. Retrieved from <http://epaa.asu.edu/ojs/article/view/1160>
- Rebell, M. A., & Block, A. R. (1982). *Educational policy making and the courts: An empirical study of judicial activism*. Philadelphia, PA: Temple University Press.
- Rivera, L. A. (2011). Ivies, extracurriculars, and exclusion: Elite employers' use of educational credentials. *Research in Social Stratification and Mobility*, 29, 71-90.
- Rochefort, D. A., & Cobb, R. W. (Eds.). (1994). *The politics of problem definition: Shaping the policy agenda*. Lawrence: University Press of Kansas.
- Ryan, J. E. (2007). Standards, testing, and school finance litigation. *Texas Law Review*, 86, 1223-1262.
- Sabel, C. F., & Simon, W. H. (2004). Destabilization rights: How public law litigation succeeds. *Harvard Law Review*, 117, 1015-1101.
- Sawchuk, S. (2014, February 3). Teachers' job protections debated in California. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2014/02/05/20lawsuit_ep.h33.html?r=1007851114
- Severance, L. J., & Loftus, E. F. (1982). Improving the ability of jurors to comprehend and apply criminal jury instructions. *Law & Society Review*, 17, 153-197.
- Stone, D. A. (1989). Causal stories and the formation of policy agendas. *Political Science Quarterly*, 104(2), 281-300.
- Stone, D. A. (2011). *Policy paradox: The art of political decision making*. New York, NY: W.W. Norton.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedure and techniques*. Newbury Park, CA: SAGE.
- Superfine, B.M. (2009). Deciding who decides questions of at the intersection of school finance reform litigation and standards-based accountability policies. *Educational Policy*, 23(3), 480-514.
- Superfine, B. M. (2013). *Equality in education law and policy: 1954-2010*. New York, NY: Cambridge University Press.

- Superfine, B. M., & Gottlieb, J. J. (2014). Teacher evaluation and collective bargaining: The new frontier of civil rights law. *Michigan State Law Review*, 2014(3), 737-788.
- Superfine, B. M., Gottlieb, J. J., & Smylie, M. A. (2012). The expanding federal role in teacher workforce policy. *Educational Policy*, 26(1), 58-78.
- Superfine, B. M. & Thompson, A. (2016). Interest groups, the courts, and educational equality: A policy regimes approach to *Vergara v. California*. *American Educational Research Journal*, 53(3), 573-604.
- Trial Tr. *Vergara v. California*. No. BC484642 (S.C. County of Los Angeles 2014).
- Urbanski v. King complaint N.Y. sup. Ct., filed March 10, 2014.
- Vaughan, D. (1999). Boundary work: Levels of analysis, the macro-micro link, and the social control of organizations. In P. Ewick, R. Kagan, & A. Sarat (Eds.), *Social science, social policy, and the law* (pp. 291-321). New York, NY: Russell Sage Foundation.
- Vergara v. California*, No. BC484642 (S.C. County of Los Angeles 2014).
- Vergara v. California* 246 Cal. App. 4th 619 (Cal. App. 2d Dist. 2016).
- Wald, P. M. (1995). The rhetoric of results and the results of rhetoric: Judicial writings. *The University of Chicago Law Review*, 62, 1371-1419.
- Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York, NY: New Teacher Project.
- Welner, K. G. (2012). Scholars as policy actors: Research, public discourse, and the zone of judicial constraints. *American Educational Research Journal*, 49(1), 7-29.
- Welner, K. G., & Kupermintz, H. (2004). Rethinking expert testimony in education rights litigation. *Educational Evaluation and Policy Analysis*, 26, 127-142.
- Wilson, P. (2000). Deficit reduction as causal story: Strategic politics and welfare state retrenchment. *The Social Science Journal*, 37, 97-112.
- Wise, A. E. (1968). *Rich schools, poor schools: The promise of equal educational opportunity*. Chicago, IL: The University of Chicago Press.
- Witt, J. F. (2004). *The accidental republic: Crippled workingmen, destitute widows, and the remaking of American law*. Cambridge, MA: Harvard University Press.

Author Biographies

Jessica J. Gottlieb is an Assistant Professor in the Department of Curriculum and Instruction at Texas Tech University. Her work focuses on understanding how education policy can expand access to high-quality STEM learning opportunities.

Ethan L. Hutt is an Assistant Professor in the Department of Teaching and Learning, Policy and Leadership at the University of Maryland-College Park. His work examines the history, development, and use of quantitative measures of schooling.

Benjamin M. Superfine is a Professor at the University of Illinois at Chicago. His research examines large-scale school reform at the intersection of education law and policy in areas such as school finance litigation reform, teacher evaluation, and collective bargaining.