

Rankings and Reactivity: How Public Measures Recreate Social Worlds¹

Wendy Nelson Espeland
Northwestern University

Michael Sauder
University of Iowa

Recently, there has been a proliferation of measures responding to demands for accountability and transparency. Using the example of media rankings of law schools, this article argues that the methodological concept of reactivity—the idea that people change their behavior in reaction to being evaluated, observed, or measured—offers a useful lens for disclosing how these measures effect change. A framework is proposed for investigating the consequences, both intended and unintended, of public measures. The article first identifies two mechanisms, self-fulfilling prophecy and commensuration, that induce reactivity and then distinguishes patterns of effects produced by reactivity. This approach demonstrates how these increasingly fateful public measures change expectations and permeate institutions, suggesting why it is important for scholars to investigate the impact of these measures more systematically.

INTRODUCTION

In the past two decades demands for accountability, transparency, and efficiency have prompted a flood of social measures designed to evaluate the performances of individuals and organizations. Measures like the cost-

¹ The authors contributed equally to this article. We are grateful for support from the Russell Sage Foundation Fellowship Program, the Law School Admissions Council (LSAC), and the Radcliffe Institute for Advanced Study. The opinions and conclusions contained in this report are those of the authors and do not necessarily reflect the position or policy of the LSAC. We thank Rick Abel, Gabi Abend, John Braithwaite, Bruce Carruthers, Alan Czaplicki, Bryant Garth, Tim Hallett, Terry Halliday, Jack Heinz, Carol Heimer, Ryon Lancaster, Tanina Rostain, Kim Scheppele, Debra Schleef, Peter Siegelman, Mitchell Stevens, and Art Stinchcombe, audiences from Australian National University, University of California at Berkeley, Harvard University, MIT, New York Law School, New York University, Northwestern University, Princeton

benefit ratios imposed on regulators, standardized tests used to evaluate students and schools, performance measures in firms, assessments of universities, and rankings of schools, firms, and hospitals have become so pervasive, scholars now characterize them as an "audit explosion" (Power 1994), "a tsunami of accountability and transparency" (Caron and Gely 2004, p. 1553), and evidence of an emerging global "audit culture" (Strathern 2000, p. 2).

Such evaluative measures are consequential and controversial. In Great Britain assessment exercises initiated during the 1980s have reorganized higher education, dictating budgets and hiring, reinventing schools as "financial bodies," and generating a "whole host of new institutions and practices" (Strathern 1997, p. 309; Shore and Wright 2000, p. 61). The growing use of quantitative indicators has transformed the meaning of accountability (Hoskins 1996). Where accountability once included many different practices, making institutions accountable now usually means making them "auditable," which often involves devising indicators to measure performance (Power 1997; Strathern 1996, 2000).

Proponents see these measures as making important information more accessible to consumers and clients, motivating organizations to improve, providing crucial feedback about policies, and extending market discipline to other institutions. Critics point to harmful, unintended consequences that occur when professionals lose discretion and actors "play to the test," focusing on indicators rather than on the qualities that the measures are designed to evaluate. Since New York began publicly ranking doctors' performance with criteria including mortality rates, for example, 79% of cardiologists surveyed report that their decision to perform angioplasty or to intervene in critically ill patients was influenced by potential effects on their scorecards (Narins et al. 2005).

Because people are reflexive beings who continually monitor and interpret the world and adjust their actions accordingly, measures are *reactive*. Measures elicit responses from people who intervene in the objects they measure. Understanding the character and consequences of this reflexivity is important. The proliferation of quantitative measures of performance is a significant social trend that is fundamental to accountability and governance; it can initiate sweeping changes in status systems, work relations, and the reproduction of inequality. Given their proliferation and their capacity to foster broad changes, both intended and unintended, in

University, and the University of Wisconsin Law School, and the *AJS* reviewers for their helpful suggestions. Ana Campos, Olivia DeBree, Corey Fields, Cassandra Malik, Cecily Harwick, and Sasha Tuzel provided expert research assistance. We gratefully acknowledge those who participated in our research, whom we promised not to name. Direct correspondence to Michael Sauder, Department of Sociology, University of Iowa, W519 Seashore Hall, Iowa City, Iowa 52245. E-mail: michael-sauder@uiowa.edu

Rankings and Reactivity

the activities that they monitor, social measures deserve closer scholarly attention.

In sociology, reactivity is usually depicted as a methodological problem. Campbell's (1957, p. 298) classic statement defines a reactive measure as one that "modifies the phenomenon under study, which changes the very thing that one is trying to measure." Reactivity blurs the distinction between the act of measuring and its object which "contaminates" results. Ethnographers express a similar concern when they advise researchers to minimize the effects of their presence on subjects' behavior so as to "avoid 'distorting' the worlds that they study" (Katz 1983). Other ethnographic approaches (cf. Burawoy 1998) regard reactivity as an inevitable part of research and urge scholars to theorize reflexivity and reactivity rather than try to eliminate it.

We agree that reactivity is an inevitable result of human reflexivity, and that social scientists should investigate reactivity more systematically. To expand our knowledge of reactivity, we propose a framework for investigating the reactivity of social measures. The core issues guiding our framework are (1) which mechanisms shape the reactivity of public measures, and (2) how to characterize the effects of the reactivity of public measures. We develop our framework for analyzing reactivity and illustrate its value through an investigation of one powerful measure, the *U.S. News and World Report* (hereafter, *USN*) rankings of law schools. Our case study of law school rankings demonstrates the distinctive characteristics of particular measures in specific locations—media rankings of law schools—but it also provides a set of concepts and questions that are important for investigating similarities and variation in the reactivity of other measures.

We first sketch the rise of social statistics and their use to describe states and evaluate performances. We next explain why reactivity remains a valuable concept for analyzing the reflexive interactions between people and measures. After briefly describing the history of rankings and how *USN* calculates law school rankings, we explain how two powerful mechanisms, self-fulfilling prophecy and commensuration, generate reactivity. We then identify three important effects of reactivity: the redistribution of resources, redefinition of work, and proliferation of gaming strategies. We conclude by summarizing our framework, suggesting its usefulness for analyzing the reactivity of other measures, and highlighting the moral implications of quantitative accountability.

ACCOUNTABILITY AND THE PROLIFERATION OF SOCIAL MEASURES

Awash in numbers, we now take social statistics for granted so much that we forget how hard it is to make them and how revolutionary their effects can be. Important historical work has illuminated how numbers have transformed the people they describe (cf. Kula 1986; Hacking 1990; Porter 1995; Carson, in press). Social numbers, of course, are nothing new. Rulers and church leaders have long tried to calculate how many men are eligible for conscription or the volume of crops or acreage subject to taxation or tithe. But such numbers were often treated as state or church secrets, the privileged information of elites that, if in the wrong hands, were dangerous (Porter 1986, pp. 18–39; Alonso and Starr 1987, pp. 11–13; Hacking 1990, p. 20). What is relatively new is the public nature of social statistics.

Use of statistics is closely linked to advances in probability and the development of nation-states and disciplines (Alonso and Starr 1987, p. 3; Porter 1986, pp. 156–57). In the 19th century, states were expanding, centralizing, and responding to new demands for public services. States needed better information about their empires. The result was an “avalanche” of numbers (Hacking 1990, pp. 1–10) that “profoundly transformed what we choose to do, who we try to be, and what we think of ourselves.” The spread of numeracy, technical advances, and the increasing legitimacy of quantification all contributed to this trend, but sometimes societies had to change dramatically before states could use these advances. Porter (1995, pp. 35–36, 45–46) describes how, initially, the weak French state was unable to conduct a proper census. Only after the state became more centralized, its administrators more numerous and better trained, and its citizens disciplined enough to permit enumeration, could it do so. Eventually, what it meant to be a state was linked to producing public statistics describing its characteristics and people (Ventresca 1995). Prewitt (1987, pp. 261–74) argues that official statistics are closely connected to democratic governance.

For Weber (1978, p. 975), “the peculiarity of modern culture,” especially its technical prowess, derives largely from its “calculability.”² Quantification, with its capacity to enhance predictability, coordination, and impersonality, is the hallmark of bureaucratic authority and helps explain why bureaucracies outperform other institutions. A shift in emphasis from quantification employed to describe relations and make predictions, to quantification used to judge and control these relations, is a small but fateful one (Foucault 1977, pp. 183–84). Discipline that constructs compliant, self-policing subjects hinges on “normalization” that defines nor-

² Urla (1993, p. 818) contends, “There are probably few features more characteristic of modernity than the notion that we can know ourselves through numbers.”

Rankings and Reactivity

mal, marks deviance, and creates the experts who maintain the boundaries.

Accountability directed toward controlling human performance through quantification and written assessments first emerged in education around the beginning of the 17th century. Not until the 19th century were these practices incorporated into the workplace (Hoskins 1996). Quantitative authority and its link to accountability and evaluation are now so secure, so bound up with being modern, that we have trouble imagining other forms of coordination and discipline or other means of creating transparency and accountability.

The long use of social statistics should not detract from the importance of contemporary efforts to extend and transform accountability. Rankings are just one manifestation of a vast proliferation of quantitative performance measures since the late 1970s (Power 1994; Strathern 2000; Hoffer 2000; Shore and Wright 2000). The calculative techniques used in business and finance are being exported to many different institutional domains. This contemporary “accountability movement” has produced diverse and important effects, and its origins are complex. According to Porter (1995), the spread of quantification in decision making is linked to efforts to mitigate conflict, overcome distrust, and coordinate across both social and physical distance. Others tie accountability and auditing to advances in neoliberal goals of shrinking government and supplanting social welfare policies with market-based policies, especially in the United States and the United Kingdom.³ Efforts to integrate economies or governments have also promulgated auditing as a way to enact transparency and accountability, as have international organizations like the United Nations, World Bank, and International Monetary Fund (Halliday and Carruthers 2005).

Rankings are one instance of widespread efforts to control public institutions and make them more accessible to outsiders. *USN*'s goal in creating rankings was not to transform law schools or “hold them accountable.” Rather, they wished to provide accessible information to educational consumers (Morse 2005), as well as raise revenue and visibility for the magazine. Editors were surprised by the furor their rankings generated. But, as we will show, we need to scrutinize the response to rankings carefully to understand the important institutional changes that performance measures set in motion.

³ Assessment tactics were common in command economies, too. Production quotas monitored productivity in the USSR and were often manipulated (Berliner 1956).

Reactivity

The concept of reactivity began as a methodological concern about the contaminating influence of measurements on their target object, especially if that object is a person (Campbell 1957; Reiss 1971; Webb et al. 1981; Katz 1983). Understood as challenging the validity of findings, reactivity has generated sustained discussions among experimentalists (Orne 1962; Campbell and Stanley 1963; Tedeschi 1981), pollsters, and survey researchers (Converse and Schuman 1974; Suchman and Jordan 1990), as well as qualitative researchers (Reiss 1971; Kazdin 1982; McCall 1984). Although definitions of reactivity vary across approaches, the basic idea is the same: individuals alter their behavior in reaction to being evaluated, observed, or measured.

Some scholars investigate reactivity in substantive rather than methodological terms. Heimer (1985), for example, analyzes reactivity as a generalized problem within the insurance industry, where price is the relevant measure. Insurance changes people's incentives to be prudent, which changes risk. "Reactive risks" are harder to manage than fixed risks, so insurers devise distinctive strategies to control them. Reactivity, for Heimer, is a continuous variable, meaning that people's responses to changing insurance incentives encompass a range of possibilities and that identifying the conditions that produce variability (e.g., institutional development, the number and type of constituencies, use over time) is key for robust explanations. Following Heimer's lead, we broaden the scope and refine the concept of reactivity by identifying patterns of reactivity and the idiosyncratic consequences attending public measures.

Reactivity is a form of reflexivity. Since reflexivity is fundamental to social life, sociology possesses a rich conceptual vocabulary for analyzing the mutuality of our responses. For example, Callon (1998) uses the idea of "performativity" to describe how economic theory shapes the economy, highlighting, as we do, the significance of self-fulfilling prophecy effects (see also Latour 1987; MacKenzie and Millo 2003; Ferraro, Pfeffer, and Sutton 2005; MacKenzie 2006). We restrict our analysis to the effects of social measures and adopt the concept of reactivity for several reasons. Most important, its connection to methodological validity makes the legitimacy of knowledge a central issue in analyzing the effects of measurement. Scholarship on knowledge production and evaluation share a fundamental concern with how particular kinds of knowing and judging are made authoritative. The best work analyzes how the authority of different forms of knowledge interact (cf. Porter 1995; Breslau 1997; Heimer and Staffen 1998; Silbey and Ewick 2003; Guetzkow, Lamont, and Mallard 2004; Carson, in press).

Casting reactivity as a threat to the validity of measures draws attention

Rankings and Reactivity

to the dynamic nature of quantitative authority and its capacity to intervene, sometimes dramatically, with other forms of authority. Instead of seeing this relationship as one that can be ameliorated with good technique, we emphasize the tension that adheres to the concept of reactivity as it mediates two understandings of measures: as valid, neutral depictions of the social world, and as vehicles for enacting accountability and inducing changes in performance. In the first view, the independence of measures and the social world they target is threatened by reactivity; in the second view, the job of measures is to harness reactivity, to create a form of reactivity that causes people to change their behavior in desired ways. This tension in seeing reactivity as either a threat or a promise is a contradiction central to the use and effects of public measures and is one that warrants analysis.

Another advantage of the concept of reactivity in capturing the reflexivity of measures is its focus on agency. The concept emphasizes how people *react* to efforts to study them, how being cast as subjects of investigation changes the behaviors of both subjects and investigators. This emphasis makes reflexivity central to the concept of reactivity, highlighting the importance of sense making as a negotiated process. People interpret measurement in diverse and evolving ways, which can change the location and expression of agency. Emphasizing *reactivity* encourages scholars to make more explicit how actors make discourse meaningful in ways that shape their behavior.

Finally, we find reactivity a useful bridging concept, one that is flexible enough to extend to interdisciplinary scholarly work on measurement, standardization, and evaluation. Reactivity has a history and a literature that resonates across fields. The value of reframing work on the impact of new measurement technologies in terms of reactivity is that it offers a familiar language and a common focus that will encourage scholars to ask similar questions and make connections across a broader range of scholarship. Doing so is crucial for understanding changes in contemporary institutions.

DATA AND METHODS

This article aims to develop a framework for analyzing the reactivity of public measures by identifying and explaining mechanisms that produce reactivity, and patterns in the effects of reactivity. A case study approach is well suited for this goal of theory elaboration because it generates detailed knowledge and sustained analysis of one type of public measure (cf. Steinmetz 2004). *USN* law school rankings were selected because, first, *USN* is the dominant ranker of law schools, virtually monopolizing

legal education.⁴ Second, unlike most fields in which only the top 25 or 50 schools are ranked, *USN* ranks every accredited law school, so the effect of rankings permeates the entire field of legal education. This density of effects enhances our efforts to map the consequences of rankings. Furthermore, pressures for accountability are increasing. All trends suggest that organizational rankings will spread and become more encompassing, so legal education is unlikely to remain exceptional. Finally, legal educators have resisted rankings more vigorously than others. Their resistance makes more explicit the assumptions behind rankings and their effects, which deepens the analysis.

The primary data for this case study are 136 open-ended interviews averaging about 45 minutes with law school administrators, faculty, and staff.⁵ These interviews permitted a broad range of respondents to describe the effects and mechanisms of reactivity instead of our presupposing them, as more structured interventions would require. Since respondents were not asked identical questions, interviews were coded for general themes rather than for precisely defined variables; consequently, the frequency of particular responses is only an approximate indication of their importance.

Two waves of interviews were conducted: one designed to capture the breadth, the other the depth of ranking effects. The first wave ($N = 66$) examined whether and how rankings affected legal education and if these effects varied across different types of schools. These interviews were conducted primarily by telephone. Our sample emphasized variations in rank, ranking trajectory, mission, location, and size. The second wave of interviews ($N = 70$) investigated how rankings permeate these institutions. We visited seven focus schools, interviewing key personnel in different organizational locations, analyzing how ranking effects interact among departments and how they are distributed within schools. Our cumulative sample includes representatives from 70 (52 current and 18 former) of the approximately 190 accredited U.S. law schools, with roughly equal num-

⁴ Other law school rankings include Brian Leiter's "Educational Quality Rankings of U.S. Law Schools," *The Gourman Report*, *Barron's Guide to Law Schools*, Brennan's Law School Rankings, MyLawSchoolRankings, and lawschool 100. No alternative ranking has the impact of *USN*. Business schools, in contrast, attend to at least five high-profile rankings, which powerfully mediates the effects of rankings (Sauder and Espeland 2006).

⁵ Interviews include 39 deans or ex-deans, 49 other administrators (e.g., deans, directors of admissions and career services, personnel in academic affairs, external relations, and development), 33 faculty members, 9 other administrators (e.g., marketing, librarians, and secretaries), and 6 leaders of national legal education associations.

Rankings and Reactivity

bers from schools in each of the four tiers.⁶ All but five interviews were tape-recorded and transcribed. We used Qualrus, a qualitative data analysis program, to organize and code our interview data. Interviews from the first wave were coded using general categories such as types or location of effects (i.e., effects on admissions, effects of quantification, changes in rank, and positive effects). We refined our coding categories based on the first wave and after the second wave recoded all interviews using these modifications.

Other data were collected or reanalyzed to corroborate and help interpret interview data. We compiled a data set from public sources on applications and yields of law schools to determine if changes in school rank affected applications and matriculation (Sauder and Lancaster 2006). We observed an admissions fair for prospective law students attended by admissions personnel from most law schools. At this event, in addition to observing interactions between administrators and students, we conducted 17 interviews, averaging about 10 minutes, with law school admissions staff, as well as 92 short nonrandom interviews, averaging about five minutes, with prospective students. We analyzed organizational material produced by law schools and public information concerning these schools (e.g., Web sites, marketing materials, press releases, and news stories) to learn how rankings affect marketing. We monitored bulletin boards and chat rooms dedicated to prospective law students for an academic year. We conducted 30 in-depth interviews with business school administrators to determine how ranking effects vary across educational fields.⁷ Triangulating sources deepens our understanding and helps mitigate the limitations of particular sources.

LAW SCHOOLS RANKINGS: HISTORY AND METHODOLOGY

Universities have been ranked for at least a century, but until the late 1980s these evaluations were almost exclusively prepared for and used by academics and administrators (Webster 1992). Beginning with *USN*'s rankings of colleges in 1983, graduate and professional schools in 1987, and *Business Week*'s ranking of business schools in 1988, popular media began

⁶ Interviews included 21 respondents from fourth tier schools, 26 from third tier schools, 22 from second tier schools, and 18 from first tier schools. Of the first tier schools, 12 schools from the top 25 and 5 schools from the top 10 were represented.

⁷ Data from two other projects were reanalyzed. One project included fieldwork and in-depth interviews with 77 law and business students (Schleef 2005). The other involved 18 months of fieldwork in an undergraduate admissions office at a liberal arts college (Stevens, in press). We conducted interviews with 12 current and prospective law students and two prelaw advisors. We also interviewed a former editor and current staff member of *USN*.

producing rankings of colleges and graduate programs created for *consumers* rather than insiders. These rankings quickly became lucrative enterprises, providing new information to prospective students and others who wanted measures of “organizational success” (Caron and Gely 2004, p. 1509).⁸

USN first ranked law schools in 1987 and ranked them annually beginning in 1990. Law schools were first divided into four tiers: the top tier listed the fifty highest-rated programs in rank order, with the remaining schools separated into the second, third, and fourth tiers and listed alphabetically within tiers. Beginning in 2004, *USN* reported the top 100 law schools by rank and divided the remaining schools into the third and fourth tiers in an alphabetical list.

The *USN* ranking is composed of four general categories: reputation, selectivity, placement, and faculty resources. Small changes have been made in rankings but their basic structure has remained the same. Reputation decides 40% of a school’s overall score followed by selectivity (25%), placement (20%), and faculty resources (15%). To compile the ranking, each school’s score is standardized. These scores are then weighted, totaled, and rescaled so that the top school receives a score of 100 and other schools receive a percentage of the top score (*USN* 2005, p. 50).

MECHANISMS OF REACTIVITY

One feature of our framework for analyzing the reactivity of measures is to specify mechanisms of reactivity. Instead of simply identifying relationships between variables, conditions, or events, a mechanism describes causal patterns that generate and sustain these relationships. A mechanism is like a causal switch that, if turned on, directs action in one way rather than another.⁹ Analyses of mechanisms produce deeper causal knowledge of social relationships and their dynamism (cf. Merton 1967, 1968; Elster 1989; Hedstrom and Swedberg 1998; Stinchcombe 2005). Explaining *how* rankings are reactive, *how* they produce the changes that they produce, helps us better understand why these measures create such powerful effects.

Most generally, rankings are reactive because they change how people make sense of situations; rankings offer a generalized account for inter-

⁸ For more on *USN* rankings methodology, see Sauder and Espeland (2006). Mel Elfin, the *USN* editor who helped launch the rankings said, “It became, essentially, our franchise” (Parloff 1998). In addition to the two regular issues featuring rankings, *USN* publishes two annual guidebooks that summarize rankings for colleges and graduate programs.

⁹ We thank an anonymous reviewer for this metaphor.

Rankings and Reactivity

preparing behavior and justifying decisions within law schools, and help organize the “stock of knowledge” that participants routinely use (Schutz 1970, pp. 137–45). Administrators consider rankings when they define goals, assess progress, evaluate peers, admit students, recruit faculty, distribute scholarships, conduct placement surveys, adopt new programs, and create budgets. One dean expressed a common sentiment: “[Rankings] are always in the back of everybody’s head. With every issue that comes up, we have to ask, ‘How is this impacting our ranking?’”

We conceptualize mechanisms of reactivity as patterns that shape how people make sense of things. Patterns emerge in the forms of feedback that rankings elicit, how attention is distributed, and the interpretive scripts people adopt. These patterns influence how people interact with rankings, each other, and other institutions. Two of the most important patterns in sense making that emerged from our analysis are self-fulfilling prophecy and commensuration. The mechanisms of self-fulfilling prophecy and commensuration produce varied forms of reactivity and, as we will demonstrate, produce some of the most significant consequences of rankings.

Self-Fulfilling Prophecies

For Merton (1968, pp. 182–83, 477, original emphasis), a self-fulfilling prophecy is a “dynamic social mechanism” and an “unintended consequence”; it is a “confident error [that] generates its own spurious confirmation,” “a *false* definition of the situation evoking a new behavior which makes the originally false definition of the situation come *true*.” Merton does not use the term “reactivity” but describes its significance. He suggests that self-fulfilling prophecies create unique methodological problems for social science because subjects change their behavior in response to prediction; prediction, then, becomes a “*new and dynamic* factor” that changes the conditions under which it comes true, which harkens back to W. I. Thomas’s famous dictum, “If men define situations as real, they are real in their consequences” (Thomas and Thomas 1928, p. 572).

We do not restrict self-fulfilling prophecies to false beliefs, requiring only that an expectation, once defined as real, amplify or confirm its effect. We define self-fulfilling prophecies as processes by which reactions to social measures confirm the expectations or predictions that are embedded in measures or which increase the validity of the measure by encouraging behavior that conforms to it.¹⁰ Rankings create expectations about law

¹⁰ Rosenthal and Jacobson’s (1968) famous study of the effects of teachers’ expectations exemplifies how measures confirm expectations. Similarly, the creation of often arbitrary categories of people (e.g., “Hispanics”) illustrates how the validity of measures

schools, and people change their behavior accordingly. We demonstrate four ways that self-fulfilling prophecies shape the reactivity of rankings: external audiences' reaction to rankings, the influence of prior rankings on surveys that determine future rank, the use of rankings to make funding decisions within universities, and how activities within schools conform to rankings criteria.¹¹

The Effects of Rankings on External Audiences

One type of self-fulfilling prophecy created by rankings involves the precise distinctions rankings create. Although the raw scores used to construct *USN* rankings are tightly bunched, listing schools by rank magnifies these statistically insignificant differences in ways that produce real consequences for schools, since their position affects the perceptions and actions of outside audiences. The distinctions produced by rankings are increasingly important and taken for granted, along with the advantages and disadvantages associated with them. As a faculty member at a top law school put it: "[Rankings] create inequality among schools that are rather hard to distinguish. They create artificial lines that then have the danger of becoming real."

Even small differences in rank affect the number and quality of applications a school receives and its yield—the proportion of accepted students who attend that school (Sauder and Lancaster 2006). Shifts in applications and yield change the selectivity score used by *USN* to compute its ranking and in this way reinforce the original differences which were often largely a product of measurement noise.¹² These results confirm the claims of administrators who carefully analyze cross-application data that describes where students matriculate when they have been accepted by multiple schools. Nearly every admissions director interviewed reported that students' decisions correlate with rankings: if a school's rank declines, they lose students to schools to which they had not lost them in the past and vice versa. Such differences can harden over time.

increases over time (Porter 1995, pp. 41–42). As a category acquires constituencies who use it, those encompassed by the category develop shared interests and identities that make the category more meaningful and "real."

¹¹ Ferraro et al. (2005, p. 9) describe how theories become self-fulfilling, "regardless of their ability to predict and explain behavior." Focusing on how economic theory has informed management, they describe three mechanisms that produce self-fulfilling effects: redesigning institutions and organizations (including measurement practices), creating norms that presume a theory, and using language that shapes people's interpretations. Our analysis reflects each of these mechanisms, although we label them somewhat differently and highlight other aspects of self-fulfilling prophecies (see also Callon 1998; Mackenzie 2006).

¹² Monks and Ehrenberg (1999) found similar patterns for undergraduate rankings.

Rankings and Reactivity

The experience of two elite schools illustrates this point. Before rankings, school A lost about two-thirds of cross-admitted students to school B. School A, however, has consistently ranked slightly higher than school B, and, over time, students admitted to both schools changed their preferences; now the vast majority (reportedly between 80%–90%) of cross-admitted students choose to attend school A. The durability of this difference no doubt increased, changing both schools' student recruitment policies. This form of self-fulfilling prophecy most harms lower-tier schools. For one dean: "If you are in the third tier, it's going to be held against you. . . . It's harder to solicit your alumni to get more revenue, to have more scholarships or to have more resources to bump up the rankings. It's a self-fulfilling nightmare." Schools must respond not only to the reactions of prospective students, but also to other constituents such as trustees, boards of visitors, and alumni, all of whom provide financial and administrative support to the schools. These reactions are most pronounced when a school's rank drops, whether this reflects changes in school quality as measured by *USN* criteria, the improved ranking of a close competitor, or fluctuations in how rankings are calculated. This process was summarized by a dean of a third tier school:

Law school faculties and the smart administrators all go, "This is a bunch of hoey. We don't care about this." Until they drop and the board of trustees says, "Hey you're dropping; why should we give you more money?" and the board of visitors from the law school say, "Man, your school's really going to pot. You haven't changed a thing . . . big changes need to be made here." And your monetary support from alumni say, "Well, I'm not sure I want to support a school that's going in the wrong direction," and your money starts to dry up. And you think, "We have got to have the money; we can't afford to lose funding or else it *will* spiral downhill and we will be a worse law school. So keeping the numbers up is key.

The Influences of Prior Rankings on Survey Responses

A second type of self-fulfilling prophecy occurs when prior rankings shape current evaluations of law schools. *USN* uses two surveys to measure the reputation of the approximately 190 accredited law schools, one sent to academics and one to practitioners. Most deans, even those with long tenures and those active in the accrediting arm of the American Bar Association (ABA), admit that they know little about most law schools they are asked to evaluate. Practitioners, whose ties to legal education are often remote, will know even less than deans and faculty. For many deans, one feature they are likely to know about other schools is their *USN* ranking. Consequently, past judgments about a school's reputation,

embodied in rank, may influence current evaluations, which reproduce and strengthen these judgments.¹³ As two administrators put it:

Well, hell, I get the rankings and I get 184 schools to rank. I know about [this school], something about [that school], I know about [my school] obviously, and I've got some buddies here and there so I feel like I know something more about some of those schools. But beyond that, guess what, I'm basing these decisions on the rankings; it's a self-fulfilling prophecy. If partners [at a law firm] do anything to prepare themselves for this [reputational survey], they probably go out and get a copy of past *USN* reports and fill it out from that.

These perceptions are supported by ongoing research on the effects of school rankings on the reputation scores derived from surveys over the 15 years of rankings. Stake (2006) finds that reputation scores slowly become aligned with overall rank, concluding that past rankings are the strongest predictor of current reputation score.¹⁴

Distributing Resources by Ranking

A third type of self-fulfilling prophecy occurs within universities if budgets become linked to rankings. Schools affiliated with larger universities must often compete for funds with other programs. Lacking other benchmarks, some administrators use rankings as a heuristic to help allocate resources. A university president informed members of the university planning committee, which included the law professor we interviewed, that "new money" would be allocated partly based on the quality of the program as defined by its *USN* rank. An important criterion was whether or not proposed plans would move a program into a top 10 *USN* rank. If budgets are allocated by current rank or potential for moving up in rank, rankings reproduce and intensify the stratification they are designed to measure. How often budgets are directly linked to rankings is unclear; however,

¹³ Most informants believe that these effects are subtle or even subconscious. No one admitted to consulting the rankings as they filled out the survey. They did suggest that prior rankings contributed to their impressions of other schools, especially schools about which they knew little.

¹⁴ Some deans dispute this effect since they do not see immediate changes in their school's reputation following changes in rank, which is consistent with Schmalbeck's (2001) analysis. Focusing on *USN* top 50 law schools, he finds that reputations are relatively durable since rankings seem to have little effect on the perceptions of survey respondents: a drop in rank during one year does not harm a school's reputation score in the following year's survey. The effects of prior rankings may be more indirect than is captured in this study. As Stake (2006) suggests, changes often take time to develop and may be most powerful for lower ranked schools since they are less well-known nationally and rankings may substitute for other information.

Rankings and Reactivity

many deans worry about how rankings form impressions and how this might influence budgets indirectly.¹⁵

Realizing Embedded Assumptions

A final example of self-fulfilling prophecy as a mechanism of reactivity is how law schools respond to the conception of legal education that is embedded in rankings factors. Rankings create self-fulfilling prophecies by encouraging schools to become more like what rankings measure, which reinforces the validity of the measure. Rankings impose a standardized, universal definition of law schools which creates incentives for law schools to conform to that definition. LSAT scores, for example, are increasingly used to define merit, and test scores now matter more than other attributes in admissions decisions. Schools may also feel pressure to abandon distinctive missions that are not measured in rankings. One dean reported that his university has always been “an access university,” one that “gives students opportunities that they would not have had elsewhere.”

A student can have a very high GPA but a very low LSAT score. That student is not going to help us in the rankings. But let's assume for the moment that that student is a minority—that's the kind of student we like to admit and to give an opportunity to even though that's going to adversely impact our rankings, and even though that student is not going to pass the bar exam the first time around. And that is our mission. We are going to have to change our mission and change our thinking if we are going to take that student's place and give it to a student with a higher LSAT score, if we are going to be driven by the rankings. Our problem, and our challenge and our opportunity, is that we are poised at the top of a tier. So it is almost impossible to resist the temptation to move into that [higher] tier [and] then that becomes a self-perpetuating situation.

When asked if his school would still admit students with the lower LSAT scores along with the higher scoring students, he replied, “It could be that for a couple of years we don't give that student the opportunity. Or we say to that student that we used to have a place for them in the day program but now we have a place for you in the night program, which is part-time and which doesn't count as much toward the rankings.”

An even more self-conscious form of this self-fulfilling prophecy occurs when administrators adopt improved rankings as an explicit goal (see Stabile 2000). In framing aspirations using *USN* criteria, organizational

¹⁵ Many university administrators report that they use potential improvement in rank to justify claims on resources, which is confirmed by a quick survey of “strategic plans” and “annual reports.”

members work to produce an organization that conforms to and makes tangible the terms of the measures, a process that renders them increasingly valid. Another version of realizing the assumptions of measurements is when members use rankings to define or market their school to external audiences. When law schools (or central administrations) tout their rankings on their Web sites, brochures, or press releases, they are complicit in producing and disseminating identities that align with rankings, which in turn shape internal processes of identification.¹⁶ The imposed organizational identity that is embedded in rankings and members' sense of identity can interact in complex ways that sometimes reinforce the meaning and legitimacy of rankings.¹⁷

Commensuration

Commensuration is a second important mechanism that produces reactive responses to rankings. Where self-fulfilling prophecies operate primarily by changing behavior in relation to altered expectations, commensuration works mainly by transforming cognition; it changes the locus and form of attention, both creating and obscuring relations among entities. Commensuration is characterized by the transformation of qualities into quantities that share a metric, a process that is fundamental to measurement. Prices, cost-benefit ratios, survey responses, and standardized tests, like rankings, are forms of commensuration. Commensuration shapes what we pay attention to, which things are connected to other things, and how we express sameness and difference.¹⁸ Three core features of commensuration change sense making: its capacity to reduce, simplify, and integrate information; the new, precise, and all-encompassing relationships it creates among different law schools and among departments within law schools; and its capacity to elicit reflection on the validity of quantitative evaluation. The first two features are inherent in commensuration, while the last one varies.

Simplifying Information

All social forms shape attention by creating patterns of exclusion and inclusion (Simmel 1971, pp. 23–35) and by erecting or transgressing sym-

¹⁶ Given law schools' public denunciations of rankings, many commentators note the hypocrisy of this practice. Other organizations publicize their rankings, too. *USN* also ranks hospitals, and high-ranking hospitals routinely advertise their standing.

¹⁷ Elsbach and Kramer (1996) found that even for elite business schools rankings threaten members' sense of their institution's organizational identity.

¹⁸ Espeland (1998) and Espeland and Stevens (1998) provide more detailed discussions of commensuration.

Rankings and Reactivity

bolic boundaries (Lamont and Molnar 2002). Commensuration as a mechanism of sense making is notable for how rigorously it simplifies information and for how thoroughly it decontextualizes knowledge. Commensuration is a means for organizing, integrating, and eliminating information.¹⁹ It simplifies in two ways: by making irrelevant vast amounts of information, and by imposing on what remains the same form—a shared metric. Rankings reduce the disparate and variable characteristics that might influence the quality of a law school to a single number. A ranking is the culmination of many commensurative practices that create many important numbers (e.g., LSAT scores, GPAs, survey results, yields, budgets). Rankings are meant to simplify complicated information; they embody decisions that make vast amounts of other information, often qualitative knowledge that is hard to assimilate to this form, irrelevant. Respondents often criticized rankings for oversimplifying law school quality, excluding important factors like student or faculty diversity, teaching, scholarship, cost, social climate, and location. For example, a letter signed by 173 deans last year was sent to all students taking the LSAT, cautioning the students: “Several commercial enterprises promote ‘ranking’ systems that purport to reduce a wide array of information about law schools to one simple number that compares all 192 ABA-approved law schools with each other. These ranking systems are inherently flawed because none of them can take your special needs and circumstances into account when comparing law schools. . . . The factors [listed by students as] most important . . . are excluded entirely or severely undervalued by all of the numerical ranking systems.”²⁰ As attention to rankings increases, it is diverted from information that is not directly linked to rankings.

In addition to making it easier to access and process information, simplification often makes information seem more authoritative. March and Simon (1958) show how processes of simplification obscure the discretion, assumptions, and arbitrariness that infuse information, as well as absorb uncertainty and contingency. Consequently, information appears more robust and definitive than it would if presented in more complicated forms.

The constraints that commensuration imposes stem from how the common metrical relationship is constructed, as well as from the relations of relevancy and irrelevancy, visibility and invisibility, that are conferred. Difference between entities is expressed as an interval on a shared metric. Commensuration makes other categories for expressing difference irrelevant as they are subsumed into this metric. Difference expressed this

¹⁹ Thévenot (1984) describes this sacrificing of information with an eye toward future uses (e.g., coding to produce standardized variables) as an “investment in form.”

²⁰ <http://www.lsac.org/LSAC.asp?url=lsac/deans-speak-out-rankings.asp>

way is a matter of more or less, rather than kind. The difference, for example, between Yale Law School, ranked first, and the University of New Mexico Law School, ranked ninety-ninth, is 98. Distinctive characteristics not captured in the rankings are less apparent and therefore less salient. Rankings do not reveal that Yale publishes nine student-edited law journals or that the University of New Mexico trains more Native American lawyers than any other law school. These characteristics may indirectly affect a school's reputation but their relevance is highly mediated. The more we attend to schools' ranks, the less we notice these other characteristics. As Simmel (1978, pp. 443–45) observed, the more naturalized a social form becomes, the less visible and more "unreal" the things that are hard to express in that form become. In this way, rankings deflect attention from many important differences among law schools, including whether a law school is public and private, whether graduates typically become solo practitioners or work in large law firms, serve poor communities or large corporations. One dean described responses to *USN* this way: "It kind of dumbs us down. I think it dumbs down students in making selections because they just look at those numbers that reflect so little, and they give them more weight than they deserve."

The radical simplification of commensuration produces decontextualized, depersonalized numbers that are highly portable and easily made public. Numbers circulate more easily and are more easily remembered than more complicated forms of information. Three pages summarize American legal education. The capacity of numbers to circulate widely is enhanced by their abstractness, which appears to eliminate the need for particularistic knowledge in interpreting them. We assume that the meaning of numbers is universal and stable. The erasure of context and the people who make numbers is crucial for their authority. We trust that the methods that accomplished this erasure were rigorous enough to constrain the discretion, politics, and biases of those who create them.

Because we think that numbers transcend local knowledge, we suppose they do not require the same sorts of interpretation and translation as other modes of communicating. This is partly why they are so useful to us. Numbers are easy to dislodge from local contexts and reinsert in more remote contexts. Because numbers decontextualize so thoroughly, they *invite* recontextualization. Since numbers seem to retain so little of their former context, it is easy to put them to new uses in new places. Believing in the universality of numbers discourages us from noticing how generative they are: that, as numbers circulate, we recreate their meaning by building them into new contexts and reinterpreting their meaning. We can use numbers to do new things without having to acknowledge that they now mean something different. So, for example, many law firms use *USN* rankings to determine the threshold for interviews. Firms interview

Rankings and Reactivity

students with lower class rankings at highly ranked schools and, conversely, require students at lower-ranked schools to have higher class rankings. This use presumes that student quality is correlated with law school quality, independent of other measures of student quality. Used as recruiting heuristics, the meaning of rankings shifts from the use intended by *USN* (as a guide to prospective law students). This change makes rankings more important and reinforces patterns of self-fulfilling prophecy.

Commensuration Unites and Distinguishes Relations

In addition to making it easier to ignore difference that is not expressed metrically, commensuration shapes attention in other important ways. It simultaneously unifies and distinguishes the objects that it encompasses or evaluates. Commensuration *unites* objects because all entities measured bear a common relationship to each other derived from their shared metric. Instead of simply being lumped together under the general category of "law school," commensuration constructs a specific relationship between each one. With law schools, we already take their unity or coherence for granted because there has already been much cultural work that naturalizes the category, "law schools." This work includes the professional and state struggles to regulate legal education and the accreditation processes of the ABA, which impose standardized practices on law schools and law students, helping to make law schools seem alike.²¹ But however alike law schools already seem to us, commensuration unifies them further by constructing a precise hierarchical relationship among them and by making the forms of heterogeneity that exist among them less visible and therefore harder to sustain, as our example of the "access school" on the cusp of tiers illustrates.²²

The metrical relationships that rankings construct have implications for law schools as an organizational field. Formerly the field of legal education was quite internally fragmented, with a small group of elite schools on top, most of which could make claims to being the best or among the best, and a wide variety of regional or mission-centered schools filling out the field. Rankings challenge this fragmentation by reducing distinctiveness to magnitude. It becomes much harder to make status

²¹ Commensuration is most jarring and seems less like measurement and more like artifice when it encompasses things that do not already seem alike. When disparate entities are included in a cost-benefit analysis, for example, we are more mindful that the relations are being *created* among these things, and the unity that commensuration imposes seems more tenuous (Espeland 1998; Espeland and Stevens 1998).

²² Within law schools, rankings factors link the fates of different departments numerically, which makes it easy to notice progress or assign blame.

claims not supported by rankings or to sustain identities that are not linked to rankings. Only Yale can now credibly claim to be "the best" law school.

At the same time that commensuration helps establish law schools as the same, it also *distinguishes* law schools by creating a precise hierarchical relationship among every school. Each law school is assigned a specific number or tier position, a location composed of a unique configuration of other numbers. Rank establishes for each school a distinct set of relations with every other school, with the interval between schools defining it as "better" or "worse" than every other law school. Over time, commensuration establishes a specific relationship between "this year's law school" and past versions of itself. A school's history can be revised from a narrative form to a series of comparisons of static objects, like photos in a chronological series, permitting schools to track "progress" or "lapses" meticulously. Assigning precise numbers to each school creates and highlights small distinctions. Differences in rankings, even miniscule ones, are closely scrutinized. One professor described the effects of this as "horrendous. Students rely on an ordinal ranking system as if it had meaning." Many schools tell painful stories of how tiny, insignificant movement in some factor, a change that reflects "nothing real or important" in their view, caused dramatic downward mobility. One dean described how his school dropped a tier one year because one class performed slightly worse on the bar exam: "It had an enormous impact because everything's at the margins. So many schools are tied in these rankings that you make a marginal switch here or a marginal switch there, you can certainly fall one tier and on occasion you can fall two tiers in these rankings. So it's a devastating kind of problem for people." Especially for schools on the cusp of tiers, small changes matter enormously and are scrutinized meticulously. Administrators complain that the tiny differences that separate many schools take on enormous weight and that such precision is false or misleading. For one law professor, "It would be as if I graded student papers to four digits."

Small differences are closely scrutinized because commensuration makes such comparisons impossible to avoid. In translating difference into magnitude, all one notices is whether one school is higher or lower, better or worse, improving or slipping, no matter how trivial the differences. Rankings are a powerful engine for producing and reproducing hierarchy since they encourage the meticulous tracking of small differences among schools, which can become larger differences over time.

Commensuration Invites Reflection on What Numbers Represent

A final way that commensuration shapes how people make sense of situations is by encouraging people to ponder the relationship between the numbers they make and use and what these numbers are supposed to reflect. This form of reactivity addresses the purchase that numbers have on the “reality” of what they measure. *USN* requires administrators to prepare numbers which it turns into new numbers called rankings. Because these new numbers are important and dissected by many parties, the relationship or “gap” between what the numbers are supposed to represent about their target and what members believe is real about their law schools elicits explanation. In other words, rankings encourage people to think about the ontological status of numbers in relation to the ontological status of the things they measure. Some people begin to question the authority of numbers as well as the production of that authority, and it is fruitful to consider how this ontological scrutiny is distributed among groups.

Desrosières (2001) describes “four possible attitudes” about the “reality” of statistics: metrological realism, derived from scientific metrology, in which the social relationship that is measured is as real as a physical object; pragmatic accounting or “accounting realism,” derived from business accounting, in which the meaning of numbers is inseparable from the trust they inspire, trust that is predicated on the “fair” standardized practices that produce the numbers (cf. Porter 1995); “proof in use” realism, in which statistics are judged real to the extent that their use produces consistent, plausible results; and constructionism which, in contrast to the other forms of realism, understands the reality of what is measured as shaped by measurement conventions. These distinctions are useful to separate analytically, even if in practice they intermingle.

Desrosières’s emphasis is statistics, but his distinctions apply to rankings. We cannot provide a detailed analysis here so we simply note some patterns. The producers of rankings at *USN* understand them as measuring real relations, rather than constructed ones, reflecting both pragmatic and metrological attitudes. Prospective law students, who know little about how rankings are made, more often adopt a metrological realist stance. Most are uninterested in ranking methodology and simply assume that rankings measure something real about the schools. A prospective law student, for example, admitted that she took rankings “quite literally” as “objective measures of quality,” ignoring their methodology.

Many administrators think rankings misrepresent their schools in harmful ways. Their deeper knowledge of their school and the effects of rankings, and their experience in helping generate the numbers they often see as arbitrary and only superficially connected to practices at their school,

makes clear to them the limitations of straightforward connections between rankings as evaluations of their school and the “reality” of their school. One administrator explained,

The rankings don't particularly capture what a lot of schools offer. Things like commitments to diversity and missions are not taken into account. These are qualitative factors, and if someone looks at the rankings they are just going to knock out your school without even a cursory review, which is a shame. That's what irritates me. It pressures you to play the numbers game. . . . It's shameful the things that we have been forced to do because of this. We're a Catholic school and we're all about honesty. Before our placement statistics were really low because we weren't reporting people who were employed in non-legal jobs, but other people were reporting them. We were forced to do the same thing because we couldn't afford for our numbers to be that low.

Detailed knowledge like this encourages constructivist understandings of rankings. But this understanding may change over time, or may be rejected by administrators who invest in rankings as strategies for gauging their performance. The views of a dean at an elite law school, someone who sees rankings as useful measures of quality, can best be classified as “proof in use realism.” For him:

Rankings have always been there. What has changed is the way in which *USN* has tried to make them more precise and to publish them more broadly. When I was looking at law school, everyone in my cohort knew which was the best school. If you got into Yale, you went there; if you didn't get into Yale, and you got into Harvard, you went to Harvard, and so on down the list. Some schools have moved around on the list but everyone knew then what the list was. So *USN* just put onto paper what everyone was already doing anyway; they have a formula for measuring what everybody cared about. . . . [*USN* rankings] are simply more complex, with more variables in their formula.

The scrutiny that rankings invite, especially in relation to the construction and meaning of numbers, is an attitude that may extend to other numbers. For example, a former director of career services told us that rankings have “made me much more skeptical about how numbers are used. I don't take them at face value anymore now that I know how loaded they can be.”

Commensuration is a powerful, prevalent, and distinctive mechanism of reactivity because it alters the distribution of attention and relevance in legal education. It changes the terms under which people assess change and make decisions; it simplifies and reduces information, standardizes sameness and difference, and inspires members to scrutinize the meaning of numbers. Commensuration and self-fulfilling prophecy represent two

Rankings and Reactivity

distinctive mechanisms of reactivity that help us better understand how rankings transform law schools. Each prompts people to change their behavior: the latter by changing expectations, the former by changing cognition. Because commensuration changes the form of information, its impact can be immediate. Self-fulfilling processes often produce more gradual change, since it takes time for people to alter their expectations and modify their behavior accordingly. Both mechanisms of reactivity will produce varied changes over time; initial responses will be different from those coming later, after rankings have diffused more broadly or become naturalized.

Law schools' initial reactions to rankings illustrate the importance of making over-time comparisons of mechanisms of reactivity. When rankings were introduced most administrators dismissed them, associating them with guidebooks that had little impact on law schools. Administrators had to learn of rankings' increasing influence. Several deans claimed to understand their significance immediately but most ignored rankings. As one administrator said: "Now the first two or three times around, no one took it seriously. I remember one of my senior colleagues coming in and saying 'I tossed [*USN* rankings] in the garbage because it's so stupid.' And I said, 'So did I.' But we didn't know. It seemed absurd to sit here guessing at reputations." Because deans underestimated their influence, some schools first refused to provide *USN* with the information they requested. *USN* responded by "estimating" the missing data (almost always, schools report, in ways that produced worse scores than schools would have otherwise had). Once angry students and alumni vented their displeasure about their schools' rankings, administrators quickly appreciated their influence.²³ Describing reaction to a disappointing ranking, a dean recalled:

The effects were immediate hysteria. I had alumni writing me left and right. I had my board of directors asking me what had suddenly happened that [we] had suddenly [dropped in the rankings]. . . . It was an irrational response because the people writing mostly actually knew about the school. I had my student body protesting, and they're here and they know in the course of one year that nothing had happened. But they all essentially were saying, "What did you do?"

Soon schools submitted information voluntarily and began to analyze their data carefully; eventually schools became more expert in *USN* methodology, contacting *USN* to criticize criteria, clarify definitions, or lobby

²³ Media reaction to rankings amplified their importance. The annual release of rankings became "news." Local newspapers carried stories about how nearby law schools fared, especially if their rank changed.

for changes. A long-time ex-dean explained, "Eventually I realized that eyeballing wasn't enough because you had to look for the little things that create differentiations and then try to game that a little bit. And we did think about that more and more." At several schools, mathematically sophisticated faculty have gone so far as to "reverse engineer" the rankings formulas to learn how their school might improve their scores. One dean said: "We've done a lot of careful studying of the *USN* methodology to figure out what counts the most and what is perhaps the most manipulable, because those are not necessarily the same things." Reactions of students, alumni, media, and peers gradually "taught" law school administrators that rankings mattered. As this lesson spread, more people began treating improved rankings as a strategic project, and their influence widened. So, despite the deep skepticism most administrators still feel toward rankings, they have become crucial professional and organizational facts that generate enormous attention and effort.

Law schools' evolving reaction to rankings suggests several lessons for studying mechanisms of reactivity. In addition to appreciating how the institutionalization of rankings changes their effects, it is important to understand how mechanisms interact. As radical simplifications, rankings circulated more broadly and quickly, and were easier to use than other evaluations of law schools. This made their impact more immediate and pronounced, since they acquired external constituencies quickly, circumventing the professional expertise of the educators who dismissed them. Over time, law schools learned that rankings were fateful, that people made important decisions using rankings, and schools began to invest heavily in improving rankings. This reinforced rankings' impact and legitimacy and set in motion the self-fulfilling prophecies described above. Commensuration exacerbated patterns of self-fulfilling prophecies.

PATTERNS IN THE EFFECTS OF REACTIVITY

Having described two mechanisms of reactivity, we now turn to the second feature of our framework for analyzing the reactivity of public measures: identifying distinctive effects of reactivity. We describe three ways that law schools react to rankings: maximization involves changing how resources are allocated; redefinition means altering organizational scripts and procedures, including job requirements and admission policies, in conformance with ranking criteria; and manipulation entails the development of gaming strategies.

Rankings and Reactivity

Maximizing Rankings by Reallocating Resources

One way that law schools react to rankings is by redistributing resources in order to optimize their rank. This reallocation produces tangible and durable changes in law schools since money and time spent for one purpose are unavailable for other uses. The most obvious redistribution of resources occurs when schools funnel large amounts of money into activities that received little funding before rankings and which could influence ranking criteria. Two widespread examples are sharp increases in spending on merit scholarships and marketing one's law school to other law schools.

USN rankings have prompted law schools to place far greater emphasis on LSAT scores than in the past. While LSAT scores have always mattered in admissions, admissions officers and deans agree that schools now do all they can to maximize this measure. Schools use several strategies to boost their LSAT scores, but the most common is to offer merit scholarships to students with high test scores.²⁴ This has prompted big changes in budgets. A faculty member at a first-tier school recalled how his school tried to raise their average LSAT score: "We were scrapping faculty salaries, not filling lines, raising money out of graduate students through our LLM program, and all of that money was going into funding three or four hundred thousand dollars in scholarships for purely number-driven reasons. And we got the LSAT up a point or two points, and we got a dean of admissions who was, in fact, paid in part based on the numbers he brought in." This redistribution is not limited to highly ranked law schools. An admissions director described the same pattern at a fourth-tier school: "We didn't have merit scholarship awards when I left in 1996. When I came back in 2000, not only were there merit scholarships to attract students with better LSATs, but the budget for [merit scholarships] exceeded the entire financial aid budget . . . for the three years that I was there before. And that was just for one year. . . . It was phenomenally more, it was more than double." Because budgets are complex, it is often hard to know how cuts are distributed; many administrators suggest, however, that this "merit money" reduces need-based scholarships. One dean, for example, said, "I have had to shift dollars away from need-based to the merit-based. The purpose of that is to pump the numbers up to get LSAT scores up. Again, this is not nuclear physics to understand."

A second prominent example of this type of reactivity is increased spending on marketing. A school's reputation is the most influential rankings factor, worth 40% of the total score. Although it is difficult to influence

²⁴ Twenty-seven administrators reported increases in merit scholarships as an impact of rankings. Opinions varied, but most did not see this as the best use of revenue.

opinions, many law schools now invest heavily in expensive brochures touting their accomplishments, which they send to their peers and practitioners.²⁵ As one dean put it, "The biggest effect of rankings has been the enrichment of printers." Law schools receive a steady stream of these brochures all year, but they literally clog faculty mailboxes in October just before *USN* surveys are filled out. These brochures are expensive to produce and distribute because they must be sent to large numbers of potential raters in order to have a chance of influencing the reputational surveys. Many administrators say their schools spend over \$100,000 per year on this type of marketing, and estimates of annual spending ranged from the tens of thousands to over a million dollars: "Oh yeah, we've all gone into the spending of, literally, hundreds of thousands of dollars on the glossies that we paper the world with. We sent these to every law professor, every dean, every associate dean, every librarian. We send them to the people who vote for *USN*." Despite huge investments in this strategy, there is no evidence that it works. Nearly every dean we interviewed admitted that they were so inundated with these materials that they barely looked at them: "The mail that I get now from all the schools on these slick glossy things—and I just literally throw them away. My secretary doesn't even give them to me. I can't imagine how bored I'd have to be to read one of those things. But we send them out because we're afraid if we don't . . . we need to get our name out there."²⁶

Rankings have clearly changed how legal educators make decisions about how resources are allocated. Many deans described feeling conflicted between improving their ranking and improving the quality of legal education they provide. One dean described the pressure this way: "I could hire a faculty member for the amount of money I spend on [marketing]; I could support 20 students for this price; I could buy a substantial number of books for our library; all of which strikes me as what this enterprise ought to be about. . . . I could almost support an entire legal writing program. I could fund a clinic. I could do any of those things. Instead I'm putting out a magazine which goes to people who aren't interested." For many deans, trying to improve ranking indicators means not investing in other worthy projects. Although administrators may experience these pressures as contradictions, boosting rankings is still broadly understood as doing "what's best for the school" since rankings

²⁵ Thirty-four administrators discussed this effect of rankings. Twenty-five respondents raised this effect without prompting.

²⁶ We asked a law professor to save this promotional material for us. The pile approached three feet in less than a year. Since he was not an administrator, he likely received far fewer brochures than the average dean.

Rankings and Reactivity

have important effects on schools. Nevertheless, deans resent the loss of discretion this implies.

Redefining Work and Policies

A second way schools react to *USN* rankings is by redefining work in ways to optimize their rank. This occurs most frequently and dramatically in administrative jobs: career services, admissions, and the dean's office. This type of reactivity can effect enduring changes, especially if new employees are hired with different skills and backgrounds. To illustrate this form of reactivity, we highlight changes in career services work at law schools.

Career Services

Rankings encourage schools to pay more attention to placement statistics. While this sounds like a positive effect, participants do not see it that way. *USN* includes two indicators of placement: percentage employed at graduation and percentage employed nine months after graduation. To improve these factors, career services personnel now expend much greater effort on learning their graduates' employment status.²⁷ All 16 respondents working in career services described this effect. One director of career services told us that she and her staff of four spent almost six weeks each year trying to track down graduates who had not replied to their multiple requests for employment information. She admitted to "exploring the possibility of hiring private detectives to locate missing students," but rejected the idea as too costly.

While students clearly benefit from career services departments that are extremely motivated to help them find jobs, there are drawbacks to this emphasis on statistics. Personnel complain that time spent on tracking students' job status decreases time spent counseling students, networking with law firms, and conducting employment seminars. One director estimated that she spent one day every week on employment statistics. As the only career counselor for more than 600 students, she said "That's time when students can't come to me." Another director reported: "The time I spend tracking down those 33 students who don't respond, I should be spending on strengthening employer contacts, developing new programs to help students be successful in their searches. [Searching for

²⁷ The University of Texas School of Law demonstrates the stakes of failing to track employment carefully. In 1998, the school dropped from 18 to 29 in *USN* rankings because 61 recent graduates failed to return their employment surveys, and *USN* assumed that 75% of nonrespondents were unemployed (Brooks 1998).

missing graduates] is the most aggravating part of my job—but the dean wants those numbers.” This emphasis on ranking criteria has changed how career services work is experienced by employees. Most career services personnel we interviewed reported that they had entered this profession because they enjoyed advising students and helping them find satisfying jobs. This part of their job has been compromised, as pressure to maximize placement statistics has increased:

The best piece of this job typically is when you are helping people, and that’s why most people go into it. When you are helping people figure out what they want to do or helping them get a job, that is the piece of it that . . . makes up for everything else. [It’s] becoming much more statistically oriented; we’re a big school to begin with so our resources here are stretched a little bit. The more resources that we spend on *USN* and on the statistics, the less time we have for the students.²⁸

This stress on placement statistics creates a pressure like that experienced by many deans; career services personnel report feeling torn between boosting numbers and helping students find the best job for them. For one:

What do you do when you have a student that has a job offer that is not perhaps what you would want for them and they come to you for some guidance? Do you tell them to hold out a little bit longer . . . Or do you tell them to take the job and worry about it later? And I think a lot of folks at this point would tell the students just to take the job and move on to the next one. And you’re not helping the students any when you do that.

And if career services personnel do not reorient their work to boost their placement statistics, they may be fired: “We have focused more closely on placement numbers, we’re very concerned about placement numbers, and our placement director is in the process of being forced to resign right now and that is clearly related to the numbers.” By all reports, turnover rates are high. Some compare themselves to athletic coaches who must produce results quickly or lose their jobs. If people are fired for their placement statistics, their replacements will likely be hired for their ability

²⁸ Fifteen out of sixteen career service personnel we interviewed discussed this topic. Thirteen reported increased pressure from deans, faculty, or students to improve employment numbers.

Rankings and Reactivity

to improve these numbers, which further institutionalizes these changes in career services work.²⁹

This pattern is repeated in admissions, where the LSAT scores of entering classes are emphasized. Many in admissions believe rankings have, in the words of one director, “completely changed the application and admissions process.” Decisions are increasingly based on ranking considerations rather than professional judgments about which students would best fit the school. One dean of admissions said: “We are now torn between decisions that will bring a better class and a class that does better according to *USN*. . . . There was this student with a 162 LSAT score [a high score for this school] and a low GPA. I thought this guy was probably immature, a slacker . . . but I couldn’t pass on that score, so I admitted him.” This loss of decision-making authority is disturbing to admissions personnel. And, as with career services positions, admissions jobs appear to be more volatile. As admissions has become more numbers oriented, experienced admissions personnel are sometimes replaced with those who embrace this goal or have better quantitative skills. Some schools are hiring staff with MBAs instead of the conventional JD. These redefinitions of career services and admissions jobs indicate how rankings have changed the conduct of work within law schools. Boosting rankings criteria now threatens to overshadow other goals, a pattern that is echoed in the dean’s office.³⁰

Gaming Rankings

Perhaps the most direct form of reactivity is trying to “game” the numbers to maximize one’s rank. We define gaming as manipulating rules and numbers in ways that are unconnected to, or even undermine, the motivation behind them. Gaming is about managing appearances and involves efforts to improve ranking factors without improving the characteristics the factors are designed to measure. No sharp lines distinguish gaming from other strategic behavior but it is useful to distinguish gaming

²⁹ Deans experience similar pressures. The dean of the University of Houston’s law school was recently forced to resign because the school’s ranking dropped. One reporter concluded that “The popularity of the annual rankings, used by prospective students, employers and job-searching scholars to gauge an institution’s credibility, has meant that schools are funneling more resources to boost their placement. It also means that deans can expect the ranking scorecard increasingly to serve as a measure of their individual job performance” (Jones 2006, p. 6).

³⁰ Other personnel also feel rankings pressure. Some schools recruit faculty with an eye toward boosting rankings, change course offerings to boost bar passage rates, or expand marketing staff. Even administrative assistants report spending more time compiling rankings data.

analytically from other strategic behavior because it produces unique consequences.

Gaming strategies are fateful partly because both actual gaming practices and rampant anxiety about these practices produce similar changes. Because many consider gaming an illegitimate response to rankings, it is subject to gossip and secrets; this secrecy threatens the scientific legitimacy of rankings since it makes it harder to assess whether measures capture what they were intended to. Gaming also changes relations with colleagues at other institutions since it breeds distrust. Distrust speeds the diffusion of gaming strategies because it increases people's anxiety about what colleagues *might* be doing, especially among closely ranked schools. Since tiny differences separate many schools, gaming encourages a "race to the bottom." Many administrators worry that gaming threatens professional values such as honesty, responsibility, and, paradoxically, transparency. Numerous articles have chastised administrators for resorting to gaming, including one written by the president of the Association of American Law Schools (Whitman 2002; see also Wellen 2005).³¹

Respondents described a wide range of gaming strategies deployed by law schools. An early, blatant example entailed schools reporting different numbers to *USN* than to their accrediting body, the ABA. The 1995 rankings issue of *USN* lists 29 (of 177) schools that reported higher median LSAT scores to the magazine than to the ABA. Detroit Mercy and the University of Alabama, for example, had four-point differentials between their *USN* and ABA medians (*USN*, March 20, 1995). This public admonishment was effective. The following year only 13 schools were chastised for discrepant scores (*USN*, March 18, 1996).

Although less nakedly executed than inconsistent reporting, many gaming strategies involve misrepresentation. Law schools now count as employed graduates with any job, including nonlegal jobs. One director of career services recounts: "What we did not have to deal with before [rankings] was intense administrative pressure to get the numbers up . . . not to help people develop career plans or help people understand how they might be able to use their gifts. . . . Now it's like 'Can you get a job in the beauty salon painting nails until these numbers are in?' So it's something that has happened on a national level with career services people and it really has a devastating effect."

³¹ 33 respondents described gaming as a problem. Many others described particular strategies used. Only one respondent believed reports of gaming were exaggerated and relatively unimportant. Since gaming tactics are ethically questionable, interviewees were reluctant to admit to gaming at their current schools, describing instead the practices of former employers and competitors. Some tactics are rare, and some stories may be apocryphal, so we report only strategies that are corroborated with public sources or reflect the direct experience of our informants.

Rankings and Reactivity

Another strategy involves reclassifying a portion of admitted students as “part-time” or “probationary” to exclude their scores from the calculation of median LSAT scores.³² These uncounted students have lower LSAT scores or GPAs but other desirable characteristics, such as interesting backgrounds, a strong record of public service, or economic, regional, or racial diversity. An associate dean at a third-tier school reported:

I know for a fact that [school X] and [school Y] have created these artificial probation programs where students with low scores come in and they don't have to count them as part of their LSAT. That's complete nonsense. . . . They have huge numbers of people on probation and it's only in the first year, and it all disappears in the second year. And they do it by taking one fewer course in the first year and then they take a makeup course in the summer. That's the rankings. We suffer because they do that and we don't.

A related strategy is to reduce the size of incoming classes by admitting fewer students into the full-time (day) program and more students into the part-time (night) program. This tactic improves both the median LSAT scores and the student-faculty ratio. A faculty member who chaired admissions at a top-tier school explained: “It's in the nature of any well-structured game that it's subject to strategic choice and strategic influence. So, one thing that we did was to shift students from our full-time into our part-time program, and I'm aware of other schools doing that.” This school moved 30 students from its full-time to its part-time program, and is considering further cuts in class size to improve rankings, a practice which disproportionately affects lower income, older, and minority students. Other, mostly highly ranked schools cut class sizes to improve scores but instead of expanding or creating part-time programs, they solicit transfer students from lower ranked local schools, a strategy that schools losing some of their best students resent. With the exception of one elite school, all of our focus schools reported that this had happened at least once to their best students.

Another gaming strategy employed by some law schools involves pressuring faculty to take spring leaves. Following ABA guidelines, *USN* counts full-time faculty teaching in the fall term for its student-faculty ratio; a faculty member on leave in the fall harms a school's ratio. Several administrators admitted this practice and many claimed it is common (see also Whitman 2002). Given that the student-faculty ratio accounts for just 3% of a school's overall *USN* rank, this suggests how meticulously schools monitor their scores.

³² Fifteen respondents reported this strategy, and most believe its use is increasing. Other reports on the proliferation of part-time programs provide corroboration.

Some schools hire their unemployed graduates temporarily to improve their employment numbers, while others (rarely) solicit more applicants, including obviously unqualified applicants, to boost their selectivity numbers. Gaming, according to many administrators, encourages them to be more distrustful of their peers, especially in their reporting of statistics like graduate placement that are difficult to verify.

The redistribution of resources, the redefinition of work, and the adoption of gaming strategies are three prominent types of reactivity that help explain how rankings affect law schools. In each case, organizational actors alter their behavior to try to influence their rank. Our account depicts the reactivity of rankings as producing serious and mostly harmful effects. The few deans who support rankings praise them for making more information available to consumers and for creating accountability. The demand for rankings proves that many find them useful for evaluating the relative performance and stature of law schools and for simplifying decisions. Rankings have also prompted law schools and their professional associations to release more information. But, as many respondents noted, rankings are deeply flawed measures, simplification can result in poorer decisions, and the unintended consequences of rankings can be deleterious. Carefully disentangling the reactivity of measures demonstrates the range and pervasiveness of these effects, whether manifest or latent.

Because rankings shape the stratification of legal education (Sauder 2006), they affect all schools, but they do not affect all schools in the same way or to the same degree. Five key sources of variation in rankings effects include: whether schools are on the cusp of tiers, their relative position in the rankings, a dean's commitment to improving rankings, geographical location, and mission. Schools on the cusp of tiers often experience enormous pressure to improve rankings. Because rankings are relative, a school's rank can shift even if nothing changes at that school. Many schools feel they must "swim furiously to stay in place," especially if failing to do so means dropping a tier. A school's position in the rankings also influences ranking effects. While members of lower ranked schools often feel demoralized by their ranking, elite schools are not immune from rankings pressure as some assume. Elite schools may not be stigmatized, but a decline in rank can produce anxiety and vigorous efforts to improve rank.

Leadership produces variation in ranking effects, too. Deans vary in their emphasis on and strategic response to rankings. Deans hired to "improve rankings" or those who adopt higher rankings as an explicit goal, either voluntary or coercively, amplify the reactivity of rankings. Ranking effects may be diminished if a law school dean resists rankings or is not pressured by key constituents to improve them. Geographical

location also shapes reactivity. Members of regional law schools may feel less rankings pressure if there are few in-state law schools and if most students wish to practice law locally. For schools aspiring to be a “national” school, rankings matter more, especially if they occupy the same niche in the same region as competitor schools. Finally, the reactivity of rankings is mediated by a school’s mission. Schools with distinctive missions, whether this entails commitments to religious traditions, public service, or providing “access,” can sometimes resist rankings pressures despite being penalized by the rankings’ universalistic and elite-based criteria.

CONCLUSION

The expansion of public measures and their capacity to change the institutions and the people they describe suggests why it is important to explain their significance. Our framework is designed to help identify the core mechanisms and effects of reactivity. Using the example of law school rankings, we show how rankings evoke self-fulfilling prophecies that gradually transform law schools into entities that conform more closely to the criteria used to construct rankings, and we describe how commensuration changes the form and circulation of information and how people attend to it. We analyze three broad patterns in the effects of reactivity, showing how rankings prompt the redistribution of resources, the redefinition of work, and gaming.

Our framework should be treated as an orienting heuristic, an animating set of concepts and questions to help scholars identify similarities and differences in actors’ response to measures; our goal is to facilitate conversations across fields rather than to develop a general theory of reactivity. For Strathern (2000, p. 2), assessment procedures are a “distinct cultural artifact,” and appreciating their distinctiveness requires precise comparisons. Our analysis is designed to demonstrate the usefulness of this framework, and we conclude by suggesting strategies for using and developing it.

First, the distinction between mechanisms and effects is helpful for framing investigations analytically. Since the same mechanism can produce different effects and different mechanisms can produce similar effects, this distinction helps identify one source of important variation in reactivity.³³ Our framework can help scholars determine if patterns associated with rankings apply to other measures. Recent research on high-stakes testing in public schools, for instance, exhibits self-fulfilling proph-

³³ This parallels Simmel’s (1971) discussion of form and content, in which the same form can include different contents and the same content can be organized as different forms.

ecies as teachers, under pressure to raise standardized test scores, "teach to the test" (McNeil and Valenzuela 2001; Meier 2002). Comparing public rankings with private performance measures also reveals interesting variation. Executives' autonomy permits them to impose and change performance measures relatively easily, which affects their institutionalization and mitigates some of the effects described in this analysis. In contrast to media-imposed rankings, if corporate performance measures produce counterproductive gaming strategies, firms can simply revise their measures (Meyer 1994, 2002). Analyzing conditions that produce variation in the mechanisms and effects of reactivity will advance our understanding of public measures.

Investigations of measurement can also be advanced by further specifying the mechanisms and effects of reactivity described here. Commensuration is a broad mechanism of reactivity, and its specific forms may generate particular effects. For example, commodification is one form of commensuration that produces effects both comparable and disparate to the effects of commensuration described above. Like other forms of commensuration, price summarizes vast information and directs attention in particular ways. But commodification also elicits distinctive social relations, such as the fetishization of commodities.

Our framework can stimulate scholars to identify new mechanisms and effects of reactivity. Recent research suggests that measures that are adopted voluntarily have different effects than those that are imposed (Zuckerman and Sgourev 2006), so it may be fruitful to consider coercion as a separate mechanism of reactivity. Not only are rankings imposed from the outside, but they also create a prisoner's dilemma for schools since they are strictly relative and punish schools who fail to conform. The fate of schools expressed in rankings is not simply intertwined, but is zero sum: one school's success may come at the expense of many others. Investigating coercion as a mechanism of reactivity may reveal why some forms of resistance failed (e.g., boycotts), why gaming techniques spread so quickly, and why rankings generate such vigilance.

Explaining variation is one crucial way to advance our understanding of the reactivity of measures, but it is also important to refine questions and concepts, and to compare explanations across fields. The distinctive approaches of different fields can advance our understanding of the reactivity of measures when they inform one another. Science studies scholars have powerful strategies for analyzing how knowledge becomes authoritative. Organizational theorists explain how fields emerge and change. Psychologists and anthropologists, in different ways, analyze cognition and classification. Scholars of professions emphasize how expertise is expanded and defended. These fields offer unique and complementary insights for understanding the reactivity of measures.

Rankings and Reactivity

One strategy for developing theory and improving our theoretical explanations of reactivity is to consider the contradictions that measurements generate in their different uses. For example, public measures create a chronic tension between validity and accountability, between taming reactivity and producing it. How to think about this tension is an important question for scholars. Some measures are designed as incentives; their purpose is to change people's behavior. But others are supposed to be neutral descriptions of the world. The contradictions that we, as producers and consumers of public measures, invest in our measures are tricky to reconcile. As scientific depictions, measures should be objective, accurate, and nonreactive. But as incentives or techniques of accountability, measures must offer crucial feedback about people's performance and shape their aspirations. We want to harness the reactivity of measures, but here timing is everything. We want a sharp temporal break between the act of measurement and the reactivity that responds to it. We want the world to stand still for measurement and, only after, change in relation to it. Measures as a series of snapshots to which people respond are acceptable, but measures as video cameras that encourage our mugging are not. Reactivity that blurs the line between object and measurement threatens the validity of our measures. As Strathern (1996, p. 4) says, "When a measure becomes a target, it ceases to be a good measure."

Reflexivity makes the sharp distinction between the act and object of measurement hard to sustain. Because we continually interpret our social worlds, objects are unstable; they emerge and are reconstructed through measurement. Just as standardized tests create gifted and underachieving students, the census produces "Hispanics," and the LSAT defines "merit." Latour (1993) characterizes this tension broadly as pressure to erect sturdy boundaries around the social and the natural. Our perpetual, elusive, and misguided efforts to protect the purity of these categories define us as "modern." Reactivity is one form of interaction that reveals how difficult it is to maintain sharp distinctions between measures and objects of measurement, description and inscription, science and society, the social and the natural. It suggests why scientific authority that is grounded in these sorts of dualisms remains a perpetual and perennially incomplete project. The independence of the social and natural worlds is threatened by our reflexivity.

In his investigations of corporate performance measures, Meyer (1994, 2002) suggests that reactivity limits the shelf life of measures. The challenge of managing reactivity in business is less to protect the purity of the measure than to protect their usefulness as incentives. Since corporations are complex organizations, assessment indicators can never capture all important aspects of performance. Over time, people learn to manipulate measures, and their narrowness and simplification can be

organizationally self-defeating. As the value of measures declines, it becomes necessary to change them in order to capture other aspects of performance. To remain useful, measures must be a moving target. But these changes may be resisted by those who benefit from particular measures and who have invested in learning to manipulate them. The challenge for *USN* is that, as a journalistic enterprise, they are unqualified to create incentive systems to change law schools. The authority of their rankings rests, precariously, on their status as valid scientific measures and *USN*'s reputation as unbiased journalists reporting "the news" rather than creating it.

Another important set of questions for understanding reactivity involves their moral implications. We do not typically associate ethics with measurement. We have shown that public measures affect the distribution of resources, redefine statuses which can become reified and enduring, produce and reinforce inequality, and transform the language in which power presents and defends itself. Once accountability is understood quantitatively and is equated with good governance, the meanings of many core values—efficiency, improvement, accessibility, transparency, responsibility, responsiveness, and even democracy—are redefined and reinscribed in our institutions as technical rather than political accomplishments. Accountability is routinely and uncritically invoked as an obvious public good, so it is especially important that scholars conduct empirically rigorous analyses.

Assessment measures permit the easy conflation of what *is* with what *ought to be*, of what normal is in the statistical and moral sense. As Foucault (1977), Starr (1987, p. 54), Hacking (1990, p. 163), and others have argued, this conflation is insidious and rhetorically potent. Characterized as decision technologies, benchmarks, guidelines, or even "best practices," performance measures move easily from one location to another, accumulating new uses and constituents. Measures, because they standardize, simplify, and provoke comparison, become sturdy vehicles for transporting practices and values from one site to another: from one country to another, from business to education, from the academy to the military, and so on. The power attending public measures is not the obvious, self-interested power of elites, nor is it monolithic; it enters, inscribes itself, and operates differently than noisy, interested power. Reactivity is nuanced and dispersed.

The blandness of audits and assessments (cf. Porter 1995, p. 51; Strathern 1996) is rhetorically useful, since it deflects attention from their capacity to change the places and people that use them. While we suspect that research on the reactivity of measures has been hampered by its rather dreary associations with bureaucrats and statistics, the aim of the present analysis is to stimulate interest in what we think is one of the

important and challenging trends of our time—the proliferation of public measures of performance.

REFERENCES

- Alonso, William, and Paul Starr. 1987. *The Politics of Numbers*. New York: Russell Sage Foundation.
- Berliner, Joseph. 1956. "A Problem in Soviet Business Administration." *Administrative Science Quarterly* 1:86–101.
- Breslau, Daniel. 1997. "Contract Shop Epistemology: Credibility and Problem Construction in Applied Social Science." *Social Studies of Science* 27:363–94.
- Brooks, A. Phillips. 1998. "UT Law School Falls from Top 25; Officials Will Appeal U.S. News." *Austin American-Statesman*, February 20.
- Burawoy, Michael. 1998. "The Extended Case Method." *Sociological Theory* 16:4–33.
- Callon, Michel, ed. 1998. *The Laws of the Markets*. Oxford: Blackwell.
- Campbell, Donald T. 1957. "Factors Relevant to the Validity of Experiments in Social Settings." *Psychological Bulletin* 54:297–312.
- Campbell, Donald T., and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin.
- Caron, Paul L., and Rafael Gely. 2004. "What Law Schools Can Learn from Billy Beane and the Oakland Athletics." *Texas Law Review* 82:1483–1554.
- Carson, John. In press. *The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics, 1750–1940*. Princeton, N.J.: Princeton University Press.
- Converse, Jean M., and Howard Schuman. 1974. *Conversations at Random*. Ann Arbor, Mich.: Wiley.
- Desrosières, Alain. 2001. "How 'Real' Are Statistics? Four Possible Attitudes." *Social Research* 68:339–55.
- Elsbach, Kimberly D., and Roderick M. Kramer. 1996. "Members' Responses to Organizational Identity Threats: Encountering and Countering the Business Week Rankings." *Administrative Science Quarterly* 41:442–76.
- Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
- Espeland, Wendy Nelson. 1998. *The Struggle for Water: Politics, Rationality and Identity in the American Southwest*. Chicago: University of Chicago Press.
- Espeland, Wendy Nelson, and Mitchell Stevens. 1998. "Commensuration as a Social Process." *Annual Review of Sociology* 24:312–43.
- Ferraro, Fabrizio, Jeffrey Pfeffer, and Robert Sutton. 2005. "Economics Language and Assumptions: How Theories Can Become Self-Fulfilling." *Academy of Management Review* 30:8–24.
- Foucault, Michel. 1977. *Discipline and Punish: The Birth of the Prison*. London: Allen Lane.
- Guetzkow, Joseph, Michele Lamont, and Gregoire Mallard. 2004. "What Is Originality in the Humanities and the Social Sciences?" *American Sociological Review* 69: 190–212.
- Hacking, Ian. 1990. *The Taming of Chance*. Cambridge: Cambridge University Press.
- Halliday, Terence, and Bruce Carruthers. 2005. "The Recursivity of Law: Global Norm-Making and National Law-Making in the Globalization of Corporate Insolvency Regimes." Working Paper. American Bar Foundation.
- Hedstrom, Peter, and Richard Swedberg. 1998. "Social Mechanism, an Introductory Essay." Pp. 1–31 in *Social Mechanisms: An Analytic Approach to Social Theory*, edited by Peter Hedstrom and Richard Swedberg. Cambridge: Cambridge University Press.

American Journal of Sociology

- Heimer, Carol. 1985. *Reactive Risk and Rational Action: Managing Moral Hazard in Insurance Contracts*. Berkeley and Los Angeles: University of California Press.
- Heimer, Carol, and Lisa Staffen. 1998. *For the Sake of the Children*. Chicago: University of Chicago Press.
- Hoffer, Thomas. 2000. "Accountability in Education." Pp. 529–44 in *Handbook of the Sociology of Education*, edited by Maureen T. Hallinan. New York: Springer.
- Hoskins, Keith. 1996. "The 'Awful Idea of Accountability': Inscribing People into the Measurement of Objects." Pp. 265–82 in *Accountability, Power, Ethos and the Technologies of Managing*, edited by R. Munor and J. Mouritsen. London: International Thomson Business Press.
- Jones, Leigh. 2006. "Law School Deans Feel Heat from Rankings." *National Law Journal* 27:6.
- Katz, Jack. 1983. "A Theory of Qualitative Methodology: The Social System of Analytical Fieldwork." Pp. 127–48 in *Contemporary Field Research*, edited by Robert Emerson. Prospect Heights, Ill.: Waveland.
- Kazdin, A. E. 1982. "Observer Effects: Reactivity of Direct Observation." Pp. 5–19 in *Using Observers to Study Behavior*, edited by D. P. Hartmann. San Francisco: Jossey-Bass.
- Kula, Witold. 1986. *Measures and Men*. Princeton, N.J.: Princeton University Press.
- Lamont, Michele, and Virag Molnar. 2002. "The Study of Boundaries across the Social Sciences." *Annual Review of Sociology* 28:167–95.
- Latour, Bruno. 1987. *Science in Action*. Cambridge, Mass.: Harvard University Press.
- . 1993. *We Have Never Been Modern*. Cambridge, Mass.: Harvard University Press.
- MacKenzie, Donald. 2006. *An Engine, Not a Camera: How Financial Models Shape Markets*. Cambridge, Mass.: MIT Press.
- MacKenzie, Donald, and Yuval Millo. 2003. "Constructing a Market, Performing Theory: The Historical Sociology of a Financial Derivatives Exchange." *American Journal of Sociology* 109:107–45.
- March, James G., and Herbert A. Simon. 1958. *Organizations*. New York: Wiley.
- McCall, George J. 1984. "Systematic Field Observation." *Annual Review of Sociology* 10:263–82.
- McNeil, Linda, and Angela Valenzuela. 2001. "The Harmful Impact of the TAAS System of Testing in Texas: Beneath the Accountability Rhetoric." Pp. 127–50 in *Raising Standards or Raising Barriers? Inequality and High-Stakes Testing in Public Education*, edited by Gary Orfield and Mindy L. Kornhaber. New York: Century Foundation Press.
- Meier, Deborah. 2002. *In Schools We Trust: Creating Communities of Learning in an Era of Testing and Standardization*. Boston: Beacon.
- Merton, Robert K. (1948) 1968. "The Self-Fulfilling Prophecy." Pp. 475–90 in *Social Theory and Social Structure*. New York: Free Press.
- . 1967. "On Sociological Theories of the Middle Range." Pp. 39–72 in *On Theoretical Sociology*. New York: Free Press.
- Meyer, Marshall. 1994. "Measuring Performance in Economic Organizations." Pp. 556–78 in *Handbook of Economic Sociology*, edited by Neil J. Smelser and Richard Swedberg. New York: Russell Sage Foundation.
- . 2002. *Rethinking Performance Measurement: Beyond the Balanced Scorecard*. Cambridge, Mass.: Cambridge University Press.
- Monks, James, and Ronald G. Ehrenberg. 1999. "U.S. News and World Report's College Rankings: Why They Do Matter." *Change* 31:42–51.
- Morse, Robert J. 2005. "U.S. News and World Report's Law School Rankings: How and Why U.S. News Does Them." Paper presented at the annual meeting of the Law School Admissions Council, Indian Wells, Calif., June.
- Narins, Craig R., Ann M. Dozier, Frederick S. Ling, and Wojciech Zareba. 2005. "The

Rankings and Reactivity

- Influence of Public Reporting of Outcome Data on Medical Decision Making by Physicians." *Archives of Internal Medicine* 165:83-87.
- Orne, M. 1962. "On the Social Psychology of the Psychology Experiment." *American Psychologist* 17:776-83.
- Parloff, Roger. 1998. "Who's Number One? And Who's Number 52, 91, and 137?" *American Lawyer*, April.
- Porter, Theodore M. 1986. *The Rise of Statistical Thinking, 1820-1900*. Princeton, N.J.: Princeton University Press
- . 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, N.J.: Princeton University Press.
- Power, Michael. 1994. *The Audit Explosion*. London: Demos.
- . 1997. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.
- Prewitt, Kenneth. 1987. "Public Statistics and Democratic Politics." Pp. 261-74 in *The Politics of Numbers*, edited by William Alonso and Paul Starr. New York: Russell Sage Foundation.
- Reiss, A. J., Jr. 1971. "Systematic Observation of Natural Science Phenomena." Pp. 3-33 in *Sociological Methodology, 1971*, edited by H. Costner. San Francisco: Jossey-Bass.
- Rosenthal, Robert, and Lenore Jacobson. 1968. *Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development*. New York: Holt, Rinehart & Winston.
- Sauder, Michael. 2006. "Third Parties and Status Systems: How the Structures of Status Systems Matter." *Theory and Society* 35:299-321.
- Sauder, Michael, and Wendy Espeland. 2006. "Strength in Numbers? The Advantages of Multiple Rankings." *Indiana Law Journal* 81:205-27.
- Sauder, Michael, and Ryon Lancaster. 2006. "Do Rankings Matter? The Effects of U.S. News and World Report Rankings on the Admission Process of Law Schools." *Law and Society Review* 40:105-34.
- Schleef, Debra. 2005. *Managing Elites: Professional Socialization in Law and Business Schools*. Lanham, Md.: Rowan & Littlefield.
- Schmalbeck, Richard. 2001. "The Durability of Law School Reputation." *Journal of Legal Education* 48:568-90.
- Schutz, Alfred. 1970. *On Phenomenology and Social Relations*. Edited by Helmut R. Wagner. Chicago: University of Chicago Press.
- Shore, Chris, and Susan Wright. 2000. "Coercive Accountability." Pp. 57-89 in *Audit Cultures*, edited by Marilyn Strathern. London: Routledge.
- Silbey, Susan, and Patricia Ewick. 2003. "The Architecture of Authority: The Place of Law and the Space of Science." Pp. 75-109 in *The Place of Law*, edited by Austin Sarat, Lawrence Douglas, and Martha Umphrey. Ann Arbor: University of Michigan Press.
- Simmel, Georg. 1971. *On Individuality and Social Forms*. Edited by Donald N. Levine. Chicago: University of Chicago Press.
- . 1978. *The Philosophy of Money*. Translated by T. Bottomore and D. Frisby. Boston: Routledge & Kegan Paul.
- Stabile, Tom. 2000. "How to Beat U.S. News: A Law School Tries to Manufacture a Better Ranking." *National Jurist* 10 (2): 19.
- Stake, Jeffrey Evans. 2006. "The Interplay between Law School Rankings, Reputations, and Resource Allocations: Ways Rankings Mislead." *Indiana Law Journal* 82:229-70.
- Starr, Paul. 1987. "The Sociology of Official Statistics." Pp. 7-57 in *The Politics of Numbers*, edited by William Alonso and Paul Starr. New York: Russell Sage Foundation.
- Steinmetz, George. 2004. "Odious Comparisons: Incommensurability, the Case Study, and 'Small N's in Sociology.'" *Sociological Theory* 22:371-400.

American Journal of Sociology

- Stevens, Mitchell L. In press. *Creating a Class: College Admissions and the Education of Elites*. Cambridge, Mass.: Harvard University Press.
- Stinchcombe, Arthur L. 2005. *The Logic of Social Science Research*. Chicago: University of Chicago Press.
- Strathern, Marilyn. 1996. "From Improvement to Enhancement: An Anthropological Comment on the Audit Culture." *Cambridge Anthropology* 19:1-21.
- . 1997. "'Improving Ratings': Audit in the British University System." *European Review* 5:305-21.
- . 2000. *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*. New York: Routledge.
- Suchman, Lucy, and Brigitte Jordan. 1990. "Interactional Troubles in Face-to-Face Survey Interviews." *Journal of the American Statistical Association* 85:232-41.
- Tedeschi, J. T., ed. 1981. *Impression Management Theory and Social Psychological Research*. New York: Academic Press.
- Thévenot, Laurent. 1984. "Rules and Implements: Investment in Forms." *Social Science Information* 22 (1): 1-45.
- Thomas, W. I., and Dorothy Swaine Thomas. 1928. *The Children in America: Behavior Problems and Programs*. New York: Knopf.
- Urla, Jacqueline. 1993. "Cultural Politics in an Age of Statistics: Numbers, Nations and the Making of Basque Identity." *American Ethnologist* 20:818-43.
- USN (*U.S. News and World Report*). 2005. *America's Best Graduate Schools*. Washington, D.C.: U.S. News and World Report.
- Ventresca, Marc. 1995. *When States Count: Institutional and Political Dynamics in Modern Census Establishment, 1800-1993*. Ph.D. dissertation. Stanford University, Department of Sociology.
- Webb, E. J., D. T. Campbell, R. D. D. Schwartz, L. Sechrest, and J. B. Grove. 1981. *Nonreactive Measures in the Social Sciences*, 2d ed. Boston: Houghton Mifflin.
- Weber, Max. 1978. *Economy and Society*. Edited by G. Roth and C. Wittich. Berkeley: University of California Press.
- Webster, David S. 1992. "Rankings of Undergraduate Education in U.S. News and World Report and Money." *Change* 45 (March/April): 19-31.
- Wellen, Alex. 2005. "The \$8.75 Million Maneuver: How Water Bills, Temp Jobs and Backdoor Admissions Help Law Schools Scale the Rankings at U.S. News." *New York Times*, July 31.
- Whitman, Dale. 2002. "Doing the Right Thing." *The Newsletter*, April.
- Zuckerman, Ezra, and Stoyan Sgourev. 2006. "Peer Capitalism: Social Structure, Motivation, and Learning in the Market Economy." *American Journal of Sociology* 111:1327-66.