Original Article

Sociological Theory
2017, Vol. 35(3) 200–227

DOI: 10.1177/0735275117726104
st.sagepub.com
SAGE

Putting the World in Orders: Plurality in Organizational Evaluation

Christof Brandtner[1]

Abstract

Sociologists have shown that external evaluation stimulates convergent organizational behavior, yet many evaluative practices are superficial or susceptible to manipulation. When does external evaluation lead to convergence in organizational fields? Organizations regularly and increasingly experience fragmented social orders based on orthogonal notions of value, or so-called plurality. I propose that the plurality of evaluative landscapes, that is, the universe of rankings, ratings, and awards in an organizational field, compromises the potential homogenizing influence of any single evaluative practice. Plurality in the evaluative landscape weakens the causal channels through which evaluative practices influence organizational behavior. Because evaluative activities are responsive to social conditions, plurality is suggested to be highest when organizations face multiple audiences, when the meaning of value is contested, and when access to evaluation is unregulated. Neoinstitutional organizational theory and the sociology of valuation, both of which inform this article, would benefit from a more integrated account of evaluative landscapes.

Keywords

evaluation, plurality, organizational theory, evaluative landscapes, field convergence

Evaluative practices have important effects on the behavior of organizations and people, as sociologists have recently established (Adkins and Lury 2011; Lamont 2012; Zuckerman 2012). Insightful empirical work has investigated the regulatory consequences of awards, ratings, and rankings with respect to both organizational behavior (Bromley and Sharkey 2017; Chatterji and Toffel 2010; Espeland and Sauder 2016; Rossman and Schilke 2014) and consumer choice (Karpik 2010; Kovács and Sharkey 2014; Luca 2011; Salganik, Dodds, and Watts 2006). Evaluations put our world in order. The central aim in existing evaluative practices literature has been to illuminate the causal mechanisms through which evaluative practices affect organizational behavior and performance. Several social processes, ranging from reactivity to status effects to categorical legitimation, explain how evaluative practices

[1]Stanford University, Stanford, CA, USA

Corresponding Author:
Christof Brandtner, Department of Sociology, Stanford University, 450 Serra Mall, Stanford, CA 94305, USA.
Email: cbrandtner@stanford.edu

influence a multitude of organizations, from universities (e.g., Espeland and Sauder 2016) to investment banks (e.g., Podolny 1993) and securities markets (e.g., Zuckerman 1999).

But external evaluation does not always create order in organizational fields, which I define as the set of organizations that are subsumed under an evaluative regime, such as restaurants, hotels, law schools, colleges, sports clubs, or countries (DiMaggio and Powell 1983). Many evaluative practices do not spark any systematic reactions on the part of the objects of evaluation. In other cases, evaluation elicits diverse and strategically differentiated responses from organizations, instead of nurturing uniformity. While the theoretical focus on causal mechanisms has been generative, an important second question has as yet been systematically neglected: why are the effects of evaluative practices on organizations not consistent across social settings?

In this article, I discuss current approaches to studying the effects of evaluation—reactivity, status, and categorical legitimation—and their limitations in establishing *when* evaluations have outcomes for those organizations within a particular industry or domain. I argue that sociological research on rankings would benefit from accounting for the fact that many evaluations within a single organizational field are contradictory and provide nonredundant, competing information about performance. Following the burgeoning volume of work on valuation in economic sociology, I refer to the coexistence of multiple legitimate social orders as *plurality*. In a plural market regime, there can be more than one winner, and winners are determined based on different criteria or by different judges (Greenwood et al. 2011; Kornberger et al. 2015; Stark 2011). While most previous studies have examined powerful and undoubtedly consequential instances of external organizational evaluation, these were usually based on cases with one judge and consistent appraisal criteria, and disregarded cases in which multidimensional evaluative practices co-occur in a setting (Glynn, Barr, and Dacin 2000; Pontikes 2012).

Much of the past literature has dealt with specific evaluative practices, such as rankings (Espeland and Sauder 2007; Sauder and Espeland 2009; Sauder and Fine 2008), reviews (Blank 2006; Luca 2011), or awards (Goode 1978; Kovács and Sharkey 2014; Rossman and Schilke 2014). Some exceptional papers have gone beyond a particular evaluation practice to consider the interplay between different evaluative practices and how they affect organizations interdependently. To complement existing theory, and in order to understand variation in the effect of evaluations, I suggest shifting the unit of analysis. Instead of focusing on a single evaluative practice and its effects, I propose that field-level outcomes of external evaluation are determined by *evaluative landscapes*, which I define as the collectivity of evaluative practices including rankings, ratings, and awards in an organizational field.

When evaluations are based on multiple orders of worth and value, variety abounds. In this article, I contend that the plurality of evaluative landscapes moderates the pathways through which evaluations structure organizational behavior in a field. Evaluative landscapes whose criteria are aligned may lead to isomorphism in organizational behavior (DiMaggio and Powell 1983; Espeland and Sauder 2007) as well as stratification of organizations' chances of success (Podolny 1993). Inconsistency in the criteria employed in the evaluative landscape may weaken or shut down the causal channels through which rankings stimulate homogeneous behavior among organizations. My argument rests on the idea that evaluative practices are potent components of the institutional environment of organizations that materialize taken-for-granted expectations, political regulatory efforts, and shared understandings of rules and roles (Fligstein and McAdam 2012; Meyer and Scott 1983; Padgett and Powell 2012). Whether or not evaluations create isomorphic organizational fields is partially contingent on such environmental properties: who is allowed to pass judgment, how are quality and performance formally regulated, and who consumes the ratings and rankings?

This article contributes to the literature on organizational evaluation by identifying a remedy for the theoretical limitations of existing accounts of the field-level effects of organizational evaluation. Advancing the concept of plurality in evaluative landscapes, the following discussion of the conditions of convergence in response to rankings brings together various sociological studies and their insights regarding organizational responses to inconsistent information (Greenwood et al. 2011; Stark 2011). The article thus speaks (1) to organizational sociology by stipulating that a multiplicity of values moderates the interplay of organizations and their environments, (2) to neoinstitutional theory by clarifying the role of evaluative practices as an institutional force, and (3) to the sociology of valuation by outlining structural conditions and causal mechanisms of evaluative plurality.

These theoretical contributions are also of relevance to sociological theory in general. Macrosociologists commonly study complex, robust structures that resist change, and there is wide agreement that institutions, categories, and status orders cannot easily be manipulated by managerial or state intervention (Berger and Luckmann 1966; Hannan and Freeman 1989; Powell and DiMaggio 1991). External evaluations, in contrast, while cheap to produce and easy to publish, often lack transparency and are inaccessible to public accountability. A small number of individuals can alter evaluation criteria at will, and often do not disclose them to the public (Blank 2006; Karpik 2010; Espeland and Sauder 2016). If the effect of all external evaluations were as dramatic as current studies claim, the omnipresence of evaluative practices and powerful evaluators in organizational environments would be disconcerting.

HOW EVALUATION SPURS ORGANIZATIONAL FIELDS INTO MOTION

The goal of external appraisals, such as rankings, ratings, and awards, is to portray the social and economic realities of the quality of goods and services or the performance of people and organizations. Yet evaluative activities have attracted scholarly attention because rankings and ratings are not only snapshots of social reality, but they can also be engines of social action (Lamont 2012; MacKenzie 2006). In reality, evaluations can make organizations move and people dance, at times literally. Baka (2015) shows that the user-generated reviews on TripAdvisor play an important role in the valuation of places, therefore presumably influencing where hotels are built and where tourist agencies send their clients. Similarly, when the International Skating Union (ISU) introduced a novel rating system for ice skaters, it changed how athletes go about composing their choreographies. As Espeland and Lom (2016:30) explain, the ISU's judging system turned out to be "so specific that skaters are essentially forced to do certain elements if they want to score well."

Evaluations are sociologically important because they make some aspects of social life visible while playing down others—be it the overall aesthetic of ice-skating or the local charm of a tourist destination. In the following section, I outline why this observation is relevant in an organizational context, specify three mechanisms that render evaluations a social force, and explain why it is a fruitful exercise to revisit the consequences of evaluation on organizations.

Mechanisms: Convergence through Reactivity, Status, and Categorical Legitimation

Macro-organizational theorists are primarily interested in external performance measurement because evaluative practices can regulate the structure of organizational fields (Bromley and Sharkey 2017; Timmermans and Epstein 2010; Zuckerman 1999). The nature of this power is soft, so organizations are usually not forced to conform to external metrics. Still, evaluative practices have been shown to shape organizational behavior to better fit the underlying premises of the evaluation through a mix of economic incentives, cultural pressures, and structural peer effects. As "market information systems," rankings, ratings, and awards provide a basis for benchmarking and comparison among sets of organizations, which can serve as a source of validation or inspiration for strategic development (Anand and Peterson 2000; Elsbach and Kramer 1996; Simon and Knie 2013).

The propensity of evaluations to create order reflects a fundamental question in macro-organizational theory: when do organizations converge and become more similar to a cultural ideal, rather than displaying significant structural dissimilarity (DiMaggio and Powell 1983; Zuckerman 1999)? Convergent behavior is when organizational structures and practices become aligned with external ideas about proper behavior. In other words, organizations accept and internalize these criteria as relevant guidelines for decision making aimed at survival and success. For instance, the main criterion for an ideal-typical corporation is shareholder value, whereas an ideal-typical advocacy organization strives for legislative change. If all organizations adapt the same criteria, the field becomes more homogeneous and identifiable as a coherent unit with a shared social reality (Berger and Luckmann 1966).

The sociological literature offers several processes that may drive the effect of evaluative practices on organizational behavior. Three popular causal mechanisms have been put forward to explain organizations' reactions to evaluation: reactivity (Espeland and Sauder 2007; Sauder and Espeland 2009), categorical legitimation (Phillips, Turco, and Zuckerman 2013; Zuckerman 1999), and self-reifying status orders (Podolny 1993; Rossman and Schilke 2014; Sauder 2006). I introduce these three arguments before discussing their common shortcomings and the conditions under which they affect organizations differently.

Reactivity. The first argument is that being evaluated causes organizations to reflect upon what they do and how they do it. This so-called reflexivity by organizations leads to changes in behavior. A rich set of contributions to this line of research is occupied with the effect of performance metrics and rankings on educational organizations, such as schools and universities. Most notably, Espeland and Sauder (2007; Sauder and Espeland 2009) established the reactivity of *U.S. News & World Report* rankings on the structuring of law schools. The authors suggest that through self-fulfilling prophecies and the commensuration of performance, people react to being evaluated. Faced with an external critic, organizations adapt to and internalize this critic's evaluation criteria as relevant guidelines for decision making pertaining to their survival and success (Espeland and Sauder 2007, 2016; see also MacKenzie 2006). Evaluations make essential paths and values visible by providing a cognitive roadmap for social action while concealing other options (Espeland and Lom 2016).[1]

Status. A second argument is that the social ties of organizations affect their behavior; organizations higher up in the status order have more ties. It is easier for well-connected organizations to accumulate resources and to perform with excellence (Chen et al. 2012; Sauder, Lynn, and Podolny 2012). Status can help buyers to rank producers with respect to their expected quality (Podolny 1993, 2005). In the case of investment banks, high-status organizations benefit from their positions by being able to charge higher prices while operating at lower cost (Podolny 1993). This accumulative advantage is one of the reasons why status orderings are self-reproducing: high-status actors stay on top because they can produce more cheaply and sell more expensively. As Sauder (2006; Sauder et al. 2012) argues for the case of schools, rankings alter status hierarchies and motivate status-seeking behavior. An organization's evaluation results thereby serve as a status marker, which can cause an "arms race" of similar reactions. Evaluative practices create and reproduce a social fabric that gives rise to the accumulative advantage effects of status.[2]

Categorical legitimation. A third argument is that evaluation exerts cultural pressures on organizations, whereby if an organization or product cannot be compared effectively to others in the same category, it is less likely to receive attention. In an influential study of such categorical compliance, Zuckerman (1999) showed that securities markets organizations must avoid confusion about their product's identity to attract reviews by critics who specialize in the appropriate product category. Social processes generate penalties for illegitimate role performance, as with category spanners; that is, products and organizations that do not fall into what the audience sees as a legitimate category tend to be ignored (Meyer and Rowan 1977; Zuckerman 1999) or create confusion (Hsu, Hannan, and Koçak 2009).

Moving Beyond Convergence

Taken together, these three arguments suggest that rankings, ratings, and awards operate as external arbiters that motivate organizations to converge toward similar behavior. The reactivity thesis holds that evaluations can direct the attention of organizational decision makers to certain structural elements that must be optimized for organizations to succeed. The status-effect argument assumes that evaluations can create a self-reproducing social order that encourages a race for status-oriented activities. The categorical compliance thesis relies on the fact that evaluations reflect the category structures and archetypes of exemplary organizations, thereby channeling external pressure to organizations so that they will conform to the audience expectations. I suggest two challenges to this argument: the co-existence of multiple evaluative practices and diverse organizational responses to them.

Multiple evaluative practices. Practically all studies of evaluative mechanisms rest on research settings in which evaluations are prominent and tell a coherent story. The *U.S. News & World Report* ranking at the center of Espeland and Sauder's (2007) study looms large in the perception of universities and law schools; the status differences in Podolny's (1993) study of investment banks are so palpable that they are reflected in the font sizes of newspaper advertisements; and the securities analysts in Zuckerman's (1999) "categorical imperative" study are uncontested adjudicators of the market. In these cases, evaluations send clear signals about leaders and laggards, and the evaluative criteria they employ are synonymous with success. To be fair, the focus on potent evaluative practices itself gets the point across: practices of rating and ranking are powerful drivers of organizational behavior. Focusing on cases involving unambiguous evaluations and neglecting cases with multiple, competing evaluations, however, bears the danger of magnifying the isomorphic effect of rankings.

In reality, organizational evaluations fluctuate greatly in their relevance and rigor. Quantitative evaluations can be formal explicit publications or mere implicit conjecture. They come in the form of stars, ladders, benchmarks, or prestigious awards. They can be based on sophisticated data (such as rankings published by the Economist Intelligence Unit or other policy think tanks), journalistic anecdotes (such as *Monocle* rankings), personal experiences (such as Yelp or TripAdvisor), or entertaining "gut feelings" (such as BuzzFeed lists). Evaluations can be crucially important to how resources are allocated in a field (such as the allotment of paying students to law schools) and how people judge organizations' reputations (Bermiss, Zajac, and King 2013), or they can be pure infotainment without any credibility or relevance to the executive suite. In most cases, more- and less-credible reviews of quality and performance are simultaneously available. In other words, the propulsive power of evaluation may not be clear-cut in many organizational fields where no single evaluative framework is accepted.

Diverse organizational responses. A second critical observation of the sociological literature on evaluations pertains to the responses by the appraised organizations. Evaluation does not always flatten the behavioral diversity of organizations. As in-depth examinations of evaluative practices in various contexts have shown, organizations may ignore or act varyingly upon the implicit recommendations of their evaluators. There are several documented unexpected consequences of evaluation. One is that external assessments can be a source of strategic differentiation. Elsbach and Kramer (1996) found that *Business Week* rankings challenged the organizational identity of business schools by explicating their standing vis-à-vis other schools. Rather than responding in kind, these schools emphasized some qualities while downplaying others and actively looked for alternative benchmarks that painted them in a more favorable image. Consistent with the idea of resource partitioning, organizations can adopt structures that differ from the configuration of other organizations (Hannan, Pólos, and Carroll 2007; Phillips and Zuckerman 2001).

Similar to the lobbying activities of corporations, evaluations can also be an opportunity to shape the institutional forces on the organization through gaming or political negotiation (Walker and Rea 2014). In a study of the graphic visualizations of markets, Pollock and D'Adderio (2012) provide further evidence that organizational evaluation is characterized not by cold calculation but by spirited negotiation. The authors find that "magic quadrants," in which industry analysts organize and judge market participants, do not simply confine organizations within a system of boxes. Instead, the information technology firms in question accumulate specific knowledge about how to escape being pigeonholed or hire professionals who negotiate how industry analysts characterize the firm (Pollock and Williams 2016). Similarly, universities regularly game the ranking system, for instance, by reclassifying students, by urging students to take paying jobs unrelated to their professional training, or by simply reporting incorrect numbers (Espeland and Sauder 2016).

Seemingly flattening evaluations can furthermore afford organizations new identities and give rise to new qualities, for instance, when a university, government, or firm emphasizes the membership in some identity group that casts an optimal light on its ranking position (Kornberger and Carter 2010:332). In an analysis of external reviews of German universities, Gülker, Simon and Torka (2012) found that evaluations are an opportunity for collegial consultation and organizational learning. Taking on a meaning that "goes beyond a purely externally forced monitoring situation," evaluations are "far removed from a[n] impersonal, interpretation-free" procedure (Gülker et al. 2012:49). They can, therefore, create a wide range of symbolic reactions that aim at improving the organization's legitimacy or brand value ("talk") while leaving the day-to-day activities ("action") relatively untouched (Brunsson 1989; Meyer and Rowan 1977).

Rather than offering a conceptual framework of ways in which organizations can diverge from or resist their evaluations, this article seeks to extend the scope of inquiry to fragmented or otherwise inconsistent instances of external evaluation. Highlighting the importance of some evaluations without taking into account their origin, purpose, and variation across different organizational forms can be misleading. Different causal pathways—reactivity, status effects, and categorical legitimation—suggest that evaluations can effectively structure and regulate organizational fields. But these otherwise convincing theories fall short of explaining the conditions under which evaluation results not in convergent behavior but instead in active diversification. Why does external evaluation lead to conformity in some organizational fields while stimulating diversity in others?

In the remainder of this article, I offer a solution to this problem in three steps. First, I propose a shift from single evaluative practices to evaluative landscapes as the unit of analysis and suggest that the plurality, or fragmentation, of criteria in such evaluative landscapes is a crucial determinant of organizational responses to evaluation. Second, I suggest that, theoretically, there are two open questions about plurality in the evaluation of organizations: the mode and the causes of plural evaluation. Finally, I introduce several exploratory propositions about how plurality interacts with the known mechanisms of evaluation and conclude with lessons for both sociologists of organizations and of evaluation.

THE PROBLEM OF PLURALITY IN ORGANIZATIONAL EVALUATION

From Evaluative Practices to Evaluative Landscapes

External measurements proliferate across many social domains, from environmental protection and industrial relations to the performance of businesses and the transparency of charitable organizations (Bromley and Meyer 2015; Espeland and Stevens 1998). As quantitative thinking and accountability norms penetrate markets and organizations, the intent of evaluative practices becomes increasingly multivalent. Universities, for instance, are exposed to "a multitude of often conflicting internally and externally derived output expectations" (Simon and Knie 2013:405). As nonprofit organizations experience burgeoning attempts to measure and compare their social impact, no dominant standard of measurement has crystallized. Instead, as Powell et al. (2017) found, the competing interpretations from academic, commercial, and associational points of view have temporarily created a truce about what it means to do good. Another example of the lack of consistency with which many organizations are evaluated is highlighted by the reported reputations of publicly listed corporations in the United States (Brandtner and Bromley 2017; Bromley and Sharkey 2017). An analysis of corporate annual reports between 1960 and 2010 shows that the number of corporate reports mentioning rankings, ratings, and awards has increased from 15 to 39 percent, as illustrated in Figure 1. Compared to other forms of impression management, this growth is commensurate with the expansion of standards and certifications, while mentions of prestigious memberships have plateaued (Gardner and Martinko 1988; Timmermans and Epstein 2010). These evaluations stem from an increasingly large number of social domains and cover an expansive array of themes. Figure 2 shows that the average number of topics covered by different rankings, ratings, or awards in each report has increased fivefold.[3] While corporations face a clearly institutionalized yardstick of economic performance in the concept of shareholder value, they are thereby exposed to an increasingly pluralistic landscape of evaluative measures.

The analytical pivot for clarifying the conditions under which rankings, ratings, and awards have field-level outcomes is a consideration of the entirety of evaluative techniques in that field instead of examining single evaluative practices. This means shifting the level of analysis from single evaluations to evaluative landscapes. *Evaluative landscape*, in contrast to a specific ranking, rating, or award, describes the universe of relevant practices in the environment of a set of organizations that aim at evaluating and comparing these organizations' performance in a standardized way. (Lamont 2012; Miller and Power 2013; Powell et al. 2017). How relevance is defined depends on the empirical research setting, but generally speaking, an evaluative landscape encompasses all evaluative practices that elicit some kind of response within the organizational field (e.g., someone reacts to being evaluated with a press release, posts the evaluation on a website, or mentions it in an annual report).

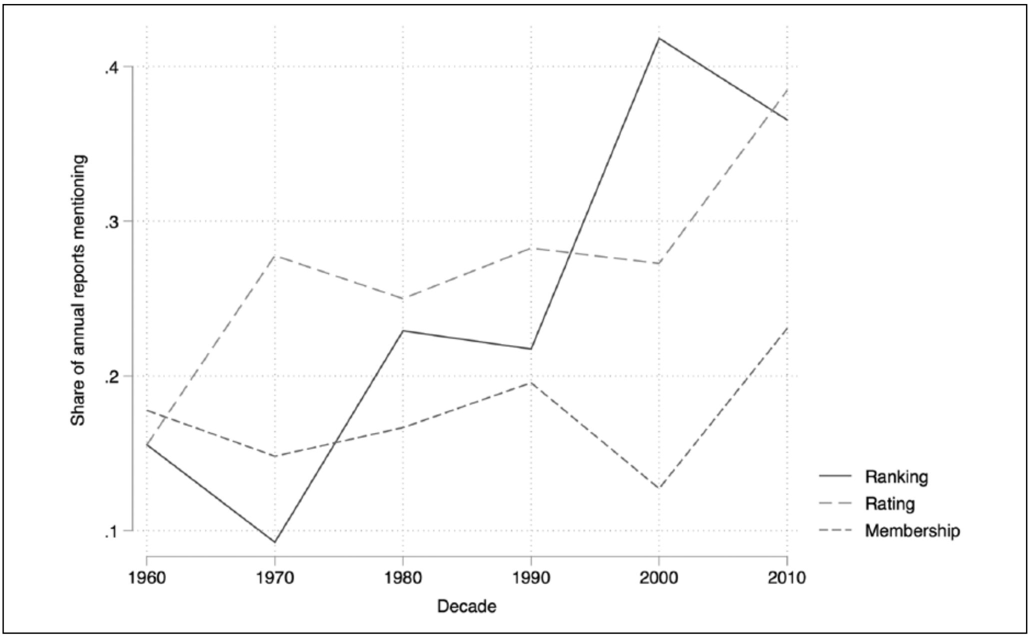

Figure 1. Prevalence of ratings, rankings, and awards in corporate annual reports.

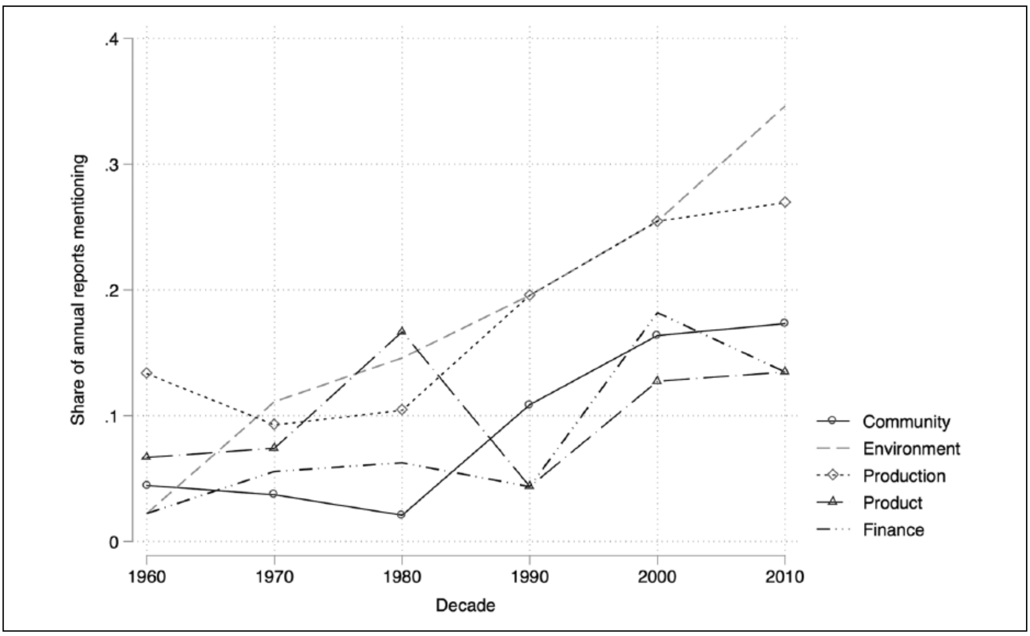

Figure 2. Content of ratings, rankings, and awards in corporate annual reports.

To further illustrate this concept, the evaluative landscape of haute cuisine restaurants includes Michelin ratings and San Pellegrino rankings, whereas the U.S. Food and Drug Administration and consumer rating sites, such as Yelp, are the only institutions evaluating regular restaurants. One example of the importance of thinking about evaluative landscapes rather than single evaluations is Christensen and Strandgaard Pedersen's (2013) discussion of the role of Michelin ratings and San Pellegrino rankings of top haute cuisine restaurants in constructing the culinary category of Nordic Cuisine, and specifically in allowing the Copenhagen restaurant Noma to be dubbed the best haute cuisine restaurant of the world. They note that the San Pellegrino top 50 rankings draw on the population of Michelin star–awarded restaurants, which requires that restaurants first attract the ranker's attention by performing well with respect to this rating. This crucial observation is possible only when both evaluative practices are included.

Plurality as a Property of Evaluative Landscapes

Taking the entirety of an evaluative landscape into account, many organizations face not a single social order but a diverse set of competing rationales (Boltanski and Thévenot 2006; Fligstein 1993; Greenwood et al. 2011; Kalberg 1980). The number of legitimate orders of evaluation varies by organizational form. For a Fortune 500 company, the predominant measure of performance is economic success, that is, shareholder value. In contrast, an academic science department can be assessed in terms of the numbers of publications or patents of its affiliated faculty and students. For a social movement organization, success can be assessed in terms of resulting legislative change, the number of supporters, or services delivered. The more legitimate roads to success that are paved with results actually measured by external evaluators, the higher is the plurality of an evaluative landscape. *Plurality* is thus defined as the presence of multiple nonredundant evaluative criteria that highlight different values and aspects of performance.

Prominent examples of reactivity to evaluative measures, such as the league tables of law schools and financial market evaluations, are cases where evaluative practices are relatively consistent and often have a significant, converging effect on organizational behavior in the evaluated organizational field. However, the same does not seem to be true for a vast set of ranked entities, whose reactivity, status-oriented action, and categorical pressures are less well understood.

City rankings provide an evocative illustration of plurality. A simple analysis of how different cities perform across various prominent rankings shows that evaluations confront city managers with an orthogonal set of expectations (see Table 1). Economic powerhouses, such as London, New York, and Hong Kong, dominate rankings of competitiveness but tend to rank poorly in livability. Conversely, medium-sized cities, such as Vienna and Copenhagen, prevail in assessments of urban quality of life. Naturally, these city league tables reflect the manifold interests and spheres of their creators and publishers. City rankings stem from consulting firms, such as Mercer and the Economist Intelligence Unit, or media outlets, such as *Forbes* and *Monocle*; others are employed as marketing tools to generate demand for a product by smart infrastructure providers, such as IBM or Siemens, or as an advocacy tactic by proponents of sustainable urban planning, such as UN Habitat (Kornberger and Carter 2010).

Urban managers refer to rankings in strategy documents, comment on their ranking performance in press releases, and celebrate their status as best-practice examples on their websites (Kornberger and Carter 2010; Lapsley, Miller, and Panozzo 2010). But despite the multiplicity of rankings on livability, sustainability, and competitiveness, city administrations pursue policies and planning strategies that are incongruent to these ideals. This example does not stand alone. While nonprofit organizations are scrutinized by watchdog platforms, such as Charity Navigator, Great Nonprofits, and GuideStar, they continue to be evaluated based on a diverse set of legitimate values (Barman 2015; Powell et al. 2017).

Table 1. Performance of Six Illustrative Cities in Rankings of Livability and Competitiveness.

	Livability			Competitiveness		
City	Mercer Quality of Living Survey	EIU Global Liveability Report	*Monocle* Most Livable Cities	At Kearney Global Cities Index	MasterCard Worldwide Centers of Commerce Index	Citi/EIU Competitive Cities Ranking
Vienna	1st	2nd	8th	13th	26th	25th
Copenhagen	9th	20th	1st	42nd	14th	23rd
Sydney	11th	6th	12th	12th	12th	15th
London	38th	54th	—	2nd	1st	2nd
New York	47th	56th	—	1st	2nd	1st
Hong Kong	70th	31st	—	5th	5th	4th
Rank out of	420	140	25	66	26	25

Note: Data from respective rankings based on 2012 data, available from the author and online.

Restaurants throughout the world are reviewed on platforms like Yelp, TripAdvisor, and Open Rice but continue to innovate, morph, and fuse in often unexpected ways (Goldberg, Hannan, and Kovács 2016; Orlikowski and Scott 2013).

The idea that organizations are exposed to multiple, inconsistent, and even orthogonal orders undermines our current understanding of how evaluations structure organizational fields. Market status theory most clearly illustrates this theoretical tension, being explicit about the premise of a uniform social order. In a seminal study, Podolny (1993) introduced the stark assumption that markets become segmented if there is uncertainty about a status order:

> The acknowledgement of a status ordering simply requires a redefinition of the market to reflect divisions on the demand side. If the market is divided, each segment can be treated as its own market and, within each segment, the significance of status can be examined. (P. 836)

In other words, Podolny assumes that there is but one status order in each social setting. This working hypothesis is important in order to support his arguments about the self-reproduction of status orders through accumulative advantage and the leakage of status through interpersonal ties. The mechanisms through which such market segmentation takes place are not explicitly stated, which puts the process into the "invisible hand" of the market (Beckert and Aspers 2011; Espeland and Stevens 1998).

The strong "status uniformity" assumption has been consistently applied in macrosociological studies of status. The idea of diverse audiences, however, calls into question the existence of unique, harmonious status orders that are valid for all buyers. In some markets, members of one audience disagree about status orders, and multiple audiences perceive different status orders (Pontikes 2012; Zhao and Zhou 2011). This status fragmentation affects the mechanisms through which rankings are assumed to reify status positions: If contractors and consumers do not share a common perception of the status order, the accumulative advantage is diminished, as the status rewards are shared by many organizations.

When rankings and ratings present organizations with a fragmented image of what counts as high performance, the unifying power of evaluation is in doubt. Before further elaborating upon these dynamics, I describe how this line of reasoning is theoretically grounded in neo-institutional theory and in the sociology of valuation.

Plurality in the Institutional Environment(s) of Organizations

That organizations face a plethora of competing expectations is a known fact to organizational theorists, even though "rather than focus on the heterogeneity of a culturally differentiated membership, scholars' emphasis has been to stress the unifying organizational principles and activities that lend cohesion and coherence to the collective" (Glynn et al. 2000:727). Scholars representing various theoretical perspectives—including the three outlined above—have noted the plurality of values in modern economies (Kornberger et al. 2015; Stark 2011).

One explicit example of this problem that is firmly rooted in North American economic sociology, but influenced by European schools of thought, is the framework of "orders of worth" (Boltanski and Thévenot 2006; Kjellberg et al. 2013; Stark 2000). Within each order of worth, evaluations called "tests" examine the value of someone's position. In other words, justification requires "qualifying (or differentiating) various objects, actors, and instruments in their environments in ways that are consistent with these logics" (Lamont 2012:208). Hence, in the case of evaluative practices, competing external evaluators may draw on different criteria and data sources to make their point.

The fragmentation of institutional orders is also the subject of intensive debate in organizational theory. Both neoinstitutionalists and organizational ecologists understand organizations as open systems that respond to audiences in their external environment, but said environment was frequently treated as a legitimacy-providing monolith that enforces a coherent set of rules and expectations (Hannan et al. 2007; Meyer and Rowan 1977; Zuckerman 1999). Recently, however, scholars from both research perspectives have noted that institutional environments are all but homogeneous (Pontikes 2012; Thornton, Ocasio, and Lounsbury 2012). Consistent with the approach to plurality outlined above, neoinstitutionalists in particular have stressed that organizations regularly face complexity resulting from competing interpretations of appropriate behavior (Dunn and Jones 2010; Greenwood et al. 2011). Organizations must carefully navigate their fragmented environments (Mair, Mayer, and Lutz 2015; Pache and Santos 2013). What is more, new types of organizations emerge at the intersection of various ecologies (Battilana and Lee 2014; Padgett and Powell 2012). The theoretical traditions underlying current theories of evaluations and rankings have, in short, moved on to the study of multivalence and heterogeneity, but our understanding of how rankings, ratings, and awards influence organizations continues to be based on assumptions of homogeneity of organizational audiences and institutional environments.

Mode and Causes of Plurality in Organizational Evaluation

Fragmented institutional frameworks pose an essential challenge to the investigation of how and which evaluations of organizations create field convergence. External arbiters facilitate the convergence of behavior among the organizations in a field for manifold reasons—be these organizations law schools, restaurants, or car manufacturers. First, evaluative practices can establish certainty about what counts as success in an organizational field, for instance, by creating industry standards or implicit norms about quality and performance. Second, the salience of legitimate categories within a field can be solidified by evaluative practices; authentic behavior is more likely when clear categories exist, such as film genres, restaurant classes, or school types. Third, through accumulative advantage and status anxiety, evaluative practices reinforce status orders in a field, which accelerates imitation among middle-status actors (Phillips and Zuckerman 2001). Convergence, however, is not always the outcome of the establishment of an evaluative regime in an industry or organizational setting.

In summary, both neoinstitutional theories of organizations and the sociology of valuation suggest that organizations frequently face multiple hierarchies, based on competing notions of value or competing logics (Greenwood et al. 2011; Lamont 2012; Stark 2011). Thus far, I have argued that this plurality of evaluative landscapes—the fragmentation of criteria used to evaluate organizations in a given field—explains why the effects of evaluation on organizations differ so dramatically. While organizational sociology, when theorizing about the effects of evaluations on organizations, usually considers single social hierarchies, many organizations face inconsistent messages about their performance, which can cause them to pick and choose information.

From this argument, two open questions particularly warrant empirical scrutiny. The first is whether the established channels through which evaluations flatten organizational worlds are in fact affected by plurality. How does plurality divert or clog the channels through which evaluative landscapes affect organizational behavior? I have also argued that the origins of status orders, league tables, and metrics are rarely scrutinized in empirical organizational research, which thereby challenges the accuracy of our understanding of why evaluations and their isomorphic effects are more prevalent in some organizations fields than in others. The second issue, therefore, is which conditions give rise to plural evaluative landscapes. To operationalize both questions, I now offer explicit propositions regarding the mode and causes of plurality in evaluative landscapes.

ON THE MODE OF PLURALITY

Plurality as a Field-level Moderator of Evaluative Effects

To restate the key proposition, the plurality of an evaluative landscape moderates the converging effect of evaluative practices on the evaluated organizations. More simply, if the criteria used by rankings, ratings, and awards are consistent with each other, evaluative landscapes may lead to convergence. If there are large gaps in what is evaluated—for instance, when rankings are based on different assumptions about the nature of performance—evaluative landscapes do not lead to field convergence.

Subjecting this proposition to empirical examination requires the measurement of the degree of consistency (or fragmentation) of the criteria of an evaluative landscape. Evaluations of some organizations are based on a single metric. For instance, academic biology departments are likely to be evaluated mainly based on the quantity of their high-quality scholarly publications, and the success of a traditional R&D department in a for-profit pharmaceutical company is measured by the number of patents. Other organizations are confronted with composite or competing indices. For instance, Padgett and Powell (2012) give an account of the historical emergence of the dedicated biotech firm. This novel organizational form emerged at the intersection of science, finance, and commerce, and is simultaneously evaluated on the basis of patents, profit, and publications. Each and every one of these dimensions can be expressed as a separate order of worth. In other words, there is variation in whether there is one dominant evaluative practice (e.g., Yelp for restaurants, *U.S. News & World Report* for law schools), there are many evaluative practices that ultimately say the same thing (e.g., San Pellegrino rankings and Michelin ratings for haute cuisine restaurants, college rankings by *Forbes* and *U.S. News & World Report*), or there is a multiplicity of dissimilar evaluative practices (e.g., the music industry, movies, cities). Plurality requires that there be disagreement regarding the criteria put forward by different rankings, ratings, and awards.

Proposition 1: The more plural an evaluative landscape, that is, the more fragmented the criteria employed by the collectivity of evaluators that assess members of an organizational field, the less organizational behavior will converge across the field.

The formal statement of this basic proposition allows for the articulation of three important caveats regarding the suggested relationship. First, the proposition implies that the expansion of the quantitative performance measurement per se, as suggested by Espeland and Stevens (1998), may in fact not lead to greater isomorphism. Each additional evaluative practice in an organizational landscape can, if its criteria contradict those in the original, weaken the isomorphic effect of rankings and create more complex organizational responses instead.

Second, the proposition highlights the causal and temporal nature of the relationship. Since rankings and ratings are primarily meant to be "snapshots" of the organizational reality, the nature of evaluations evidently follows from this snapshot. In the long run, organizational fields and their evaluative landscapes coevolve endogenously. Industry indices, such as the Dow Jones Sustainability Index, or nonprofit intermediaries, such as Charity Navigator or the Global Impact Investing Network, are examples of how organizations seek to influence the soft regulation to which they or their stakeholders are exposed (Barman 2016; Powell et al. 2017; Walker and Rea 2014). In the short run, however, changes in the cultural and social environment of organizations occur in response to rankings and ratings. Evaluative practices set the field in motion. They can reinforce and perpetuate isomorphic structures or stimulate differentiation because their status ordering has real market effects, guides consumers in choosing a legitimate offering, and causes reactivity. Changes in evaluations are therefore expected to modify incentives and redraw the cognitive maps that lead to immediate behavioral changes (Espeland and Lom 2016).

Finally, evaluations are of course not the only source of isomorphic behavior (DiMaggio and Powell 1983). Because of the "intersectionality" of different coercive, normative, and mimetic pressures, it is important to remember that not all isomorphic fields are necessarily governed by a uniform evaluative landscape and that, conversely, not all powerful rankings prompt their subjects to respond identically.

Mechanisms: Plurality Decreases Evaluative Certainty, Category Salience, and Status Anxiety

How does plurality reduce the likelihood of convergent behavior in organizational fields? That it does so follows from the insight that plurality weakens—or shuts down—the established causal channels through which evaluative practices are expected to affect organizational behavior: reflexivity, categorical legitimation, and status effects. Table 2 offers a synopsis of these causal channels and how plurality alters them.

Evaluative certainty. By quantifying the qualitative, evaluative landscapes may establish certainty about the conditions under which organizations operate. They turn incommensurate biographies into comparable cases, which thus opens them to managerial intervention (Heimer 2001). From the perspective of organizations, however, not all evaluative landscapes create certainty about what is proper behavior. Evaluations of vague or contested outcomes (such as being the "best" employee, employer, nonprofit organization, university, city, etc.) can be influential even if the concepts they intend to measure are ambiguous or contradictory. For instance, city managers are uncertain about the meaning of concepts such as sustainability or livability, so they look to evaluations for clarification. Evaluative practices and their proponents can therefore also reflect and sustain value-plural systems "by providing and diffusing alternative definitions of worth" (Lamont 2012:11).

This implies that the relationship between evaluative standards and organizational convergence depends on whether the evaluative practices can create certainty about the criteria

Table 2. Causal Channels through which Evaluative Practices Influence Organizational Behavior.

Causal Channel	Effect of Evaluation	Organizations' Motivation for Reaction	Effect of Plurality on Causal Channel
Reactivity	Commensuration and seduction	Reflexivity; following a cognitive map	Diminishes certainty about evaluative criteria
Status effects	Accumulative advantage, information revealed about social network position	Status anxiety	Reduces status anxiety, splits accumulative advantage
Categorical legitimation	Legitimacy penalty and oversight if not "in category"	Legitimacy seeking	Decreases categorical salience and contrast

of organizational effectiveness. Uniform evaluative landscapes provide organizations with certainty about "what counts" in an organizational field (Espeland and Stevens 1998; Stark 2011). More plural landscapes, however, are less likely to generate certainty, which in turn results in the absence of a clear template for appropriate behavior, as contradictory evaluative practices interfere with the process of internalizing and reflecting upon rankings, ratings, and awards (Sauder and Espeland 2009). An inconsistent evaluative landscape does not make for a guiding cognitive map (Espeland and Lom 2016).

> **Proposition 1a:** The more plural an evaluative landscape, the less it provides evaluative certainty, and the less organizational behavior will converge across the field.

Categorical salience. The absence of certainty about legitimate templates is also relevant for categorical legitimation as a process through which evaluations can prompt organizational responses (Hsu et al. 2009; Zuckerman 1999). As Kovács and Hannan (2010) demonstrate, when categories have very fuzzy boundaries, ambiguous categorical membership is less of a disadvantage. Building on data from the rating website Yelp, the authors found that both advantages of fitting into a social category—clarity on the part of the customers and focus on the part of the producer—are attenuated when categories are fuzzy.

While the categorization literature does not explicitly draw on the arguments of inconsistent institutional environments (Greenwood et al. 2011; Meyer and Scott 1983) or multiple values (Boltanski and Thévenot 2006; Stark 2011), ultimately the concept of ambiguous audience evaluations aligns with these ideas. As Pontikes (2012) found in a study of U.S. software organizations, for instance, different audiences have dissimilar viewpoints when evaluating organizations that do not fit a single category. In fact, the author showed that innovation-friendly groups, such as venture capitalists (market makers in Pontikes's terminology), prefer diverse category membership, and the flexibility associated with it, to categorical purity. This finding is in contrast with Zuckerman's (1999) investigation of security analysts, who embody the position of third-party critics or so-called market takers and who, Pontikes found, are more likely to penalize deviations from existing categories. A more fragmented set of evaluators, hence, is expected to reinforce existing category structures less than a single evaluator. Plural evaluative landscapes can be expected to decrease the salience of cultural categories and to decrease the negative effect of category spanning. As the salience of categories declines, so do the penalties for not having a clear categorical identity.

> **Proposition 1b:** The more plural an evaluative landscape, the lower is the salience of categories, and the less organizational behavior will converge across the field.

Status anxiety. The third mechanism through which plurality decreases the converging field-level effects of evaluation pertains to organizations' social standing. As discussed above, the plurality thesis poses a serious challenge to Podolny's (1993) status theory: in contexts with fragmented evaluations, the effect of accumulative advantage is lessened, because not all suppliers and customers agree on who is on top of the status order. In contrast, most empirical studies of orders of reputation or status assume that there is some kind of objective quality of the goods and services produced by organizations (Podolny 2005; Zhao and Zhou 2011). In a world with a consensus about the meaning and markers of premium quality, the effects of evaluation on organizations should not vary situationally. But for singular or unique entities (e.g., places or cultural consumer goods, such as films or books), values are often found to be incommensurable (Heimer 2001; Karpik 2010). If this is true, organizational environments in which evaluative criteria are fragmented will not produce the same status anxiety that, for instance, Sauder and Espeland (2009) described as an important driver of organizational responses to rankings. Hence, status-oriented action aimed at dominating the competition is less likely. From the point of view of status research, plurality is expected to be associated with more divergent behavior, as producers of goods try to appeal to a diversity of consumer tastes.

> **Proposition 1c:** The more plural an evaluative landscape, the less organizations will experience status anxiety, and the less organizational behavior will converge across the field.

The consideration of fragmented evaluative landscapes implies another proposition with regard to the unequal distribution of convergent behavior at the intersection of status and category. It is established knowledge that the social standing of an organization moderates its response to evaluations. For instance, Phillips and Zuckerman (2001) found that middle-status law firms and venture capitalist firms conform to categorical pressures more than both low- and high-status organizations. This "middle-status conformity" thesis also allows for an interesting derivation in relation to the moderating effect of high perceived status on organization-level analysis that can be tested empirically. Perceived status is likely to have a null effect on organizational behavior if the criteria in the evaluative landscape are inconsistent. The cumulative advantage of considering oneself high status is expected to be less significant if different suppliers and consumers perceive different status orders.

> **Proposition 1d:** The more plural an evaluative landscape, the weaker is the moderating effect of perceived organizational status on the converging pressure of evaluative standards.

Alternative Moderators: Number, Structure of Information, Salience

Previous scholars have offered a number of alternative explanations for why not all rankings, ratings, and awards lead to similar outcomes, including the number and credibility of evaluative practices, the structure of the information, and the salience of the evaluation.

Number and credibility of evaluations. One simplified explanation is that the mere number of evaluative practices in an organizational field explains their isomorphic impact on organizational behavior. Such a theory could argue that evaluations choose a common course of action only when there is a single source of information; it could also argue that the effect is strongest when there is a large number of unanimous sources of quantitative information. Without knowing the quality of the information provided by an evaluation, these hypotheses are not only contradictory but also futile. A more fruitful hypothesis is that, since evaluations are in competition with each other, the evaluative practice with the most credible and functional criteria dominates. In reality, however, the world of evaluation is usually not monopolistic, but oligopolistic. Some 50 years ago, to draw on a previously introduced example, cities were roughly compared in terms of air pollution, delinquency rates, and drug use, and there was no systematic high-profile evaluation. Today, sophisticated rankings of cities measure abstract concepts such as sustainability, innovation, or quality of living (Kornberger and Carter 2010; Rogerson 1999). A myriad of researchers, media outlets, consultancy firms, and technology companies publish rankings. They vary widely in their use of data and in their methodological rigor, and no single ranking stands out. In other words, neither of these accounts—the lack of contestation or dominance through credibility—explains why evaluations are consequential in some contexts and not in others.

Structure of information. A second alternative explanation is how evaluative practices structure information in a given field. Rossman and Schilke (2014) showed that there is a binomial distribution of returns based on a movie's Oscar appeal. The authors argued that this follows from the discontinuous status structure, that is, how performance information is presented to movie producers (Sauder 2006). The narrative for why the Oscars differ so markedly from rankings such as, in Espeland and Sauder's (2007) case, the *U.S. News & World Report* rankings, is intuitive yet hypothetical. To argue that the Oscars affect movies differently than rankings because of the structure of performance information neglects the key fact that winning Oscars is not the only indication of a movie's success. There are many more roads to success than the "expert opinion" of the Academy (such as box office appeal). In fact, information is rarely structured in only one way in an organizational field, as Orlikowski and Scott (2013) illustrated in their comparison of the American Automobile Association and TripAdvisor hotel valuations. In the movie industry, credible ratings and rankings of movies complement prestigious awards (e.g., Rotten Tomatoes, IMDb; cf. Hsu et al. 2009).

Salience of evaluation. A third explanation also supports the view that evaluative landscapes are a valuable unit of analysis for better understanding the prevalence of evaluative effects. An important dimension of variation in evaluative landscapes across fields is how visible are these evaluations. As Salganik et al. (2006) found in an online experiment, status inequality is strongly amplified by the visibility—or salience—of social evaluations. The salience of evaluation varies vastly between different fields, which raises the question as to whether audiences (consumers as well as funders) pay attention to publicly available metrics about the performance of an organization. Even more importantly, salience also varies over time: 20 years ago, attempts to measure the social impact of nonprofit organizations were rare (Barman 2007). Knowing which evaluative practices are meaningful enough for researchers to care about is difficult. For instance, one-shot rankings on social media platforms, such as BuzzFeed, are disregarded as a serious source of performance information by both the ranked organizations and their audiences, and therefore largely irrelevant. This is because evaluative practices are consequential to organizations only when these organizations believe their clients base their decisions on evaluations (Blank 2006). More salient evaluative landscapes will therefore see more convergent organizational behavior, but this observation does not invalidate the plurality argument.

ON THE CAUSES OF PLURALITY

Organizational Evaluation as Black Box

Plurality complicates the linear relationship between organizational evaluation and organizational outcomes. While sociological research has furthered our understanding of *how* evaluations affect organizations, *when* and *where* authoritative evaluations emerge remains unclear. Why does the plurality of evaluative landscapes differ across different organizational fields?

The three theoretical frameworks are somewhat unclear about the origin of social orders; in the context of status-based explanations, our knowledge of the sources of status orders in markets is limited (Gould 2002; Piazza and Castellucci 2014). Similarly, the genesis of categories is understudied (Durand and Paolella 2013; Hannan et al. 2007), as are the origins of institutions (Greenwood et al. 2008).

This lack of clarity also extends further to a scarcity of research on the emergence of organizations and markets. With some exceptions at the interstices between organization studies and social movement theory (Fligstein and McAdam 2012; Rao, Monin, and Durand 2003; Rao, Morrill, and Zald 2000) and network theory (Padgett and Powell 2012), the major intellectual streams in organizational theory have paid little attention to the macrosociological dynamics that create evaluative practices and, more generally, social orders. As such, evaluations are usually understood simply as a reflection of social conditions, without giving much attention to the processes that determine evaluative landscapes (Beckert and Aspers 2011). Other questions also remain unanswered. For instance, do changes in the criteria of evaluations merely reflect the Zeitgeist, or are there political processes involved in the collective construction of what constitutes a high reputation (Porter 1996)?

The origins of evaluations are often treated as a black box or an outgrowth of sweeping environmental processes, such as rationalization and globalization (Meyer and Bromley 2013; Porter 1996). Few studies of organizational responses to rankings, ratings, and awards consider the actual production of evaluation (i.e., the activities of the evaluators) or the influence that organizations have in structuring and shaping their evaluative environment. This is problematic, as "[p]rocesses that generate an institution are often different from those responsible for its reproduction" (Powell, Packalen, and Whittington (2012:438). Being ignorant about the origins of evaluations leads to the implicit assumption that evaluations exist because they fulfill a specific function (Goode 1978). But evaluations once they are published, can be consumed, used, and repurposed by anyone.

The motivations of rankers are often quite different from their ascribed or intended functions. Double-entry bookkeeping is a good illustration of this dichotomy. In his 1977 business history of the United States, Chandler described the practice of systematically logging expenditures and revenues as a functional solution to coordination problems between firms. The emergence of double-entry bookkeeping, as Padgett and McLean (2006) counter, is in fact the outcome of intricate network interactions and only later became the institutional foundation of capitalism. Carruthers and Espeland (1991) describe double-entry bookkeeping as having fulfilled an important rhetorical role and not just a technical purpose. As Macintosh and Quattrone (2010) further illustrate with their discussion of the Parmalat accounting scandal, today's accounting practices are not so much a machine meant to provide answers but one that raises questions and opens discussion about economic performance.

Performance metrics are thus not always meant to facilitate accountability, increase transparency, or create trust. Understanding why the effects and nature of evaluations differ across fields requires an investigation of the conditions under which evaluations are created, and specifically the people and professions that engage in evaluative activities.

The sociology of valuation offers three important insights into the problem at hand. First, in many social settings, more than one activity can be counted toward a positive evaluation, and the consequent competing social orders allow for multiple, rather than just one, justifiable actions. Second, the values and ideas in a given field are abstractions that can, under some conditions, become manifest in concrete organizational actors and social structures; evaluators and evaluative practices are manifestations of legitimated ideas. Third, evaluations produce and reflect values and make them visible; therefore, they are the basis for building reputation, status, and prestige. Who and what governs this value-producing process of evaluation—cultural institutions, legal governance, and the involved actors—should thus determine how organizations are evaluated.

Plurality as a Materialization of the Institutional Environment

A preliminary answer to the question about the causes of plurality is that organizations face fragmented criteria in their rankings, ratings, and awards because there exist competing institutional understandings of value. To elaborate several concrete propositions, I in part draw on a fruitful discussion about the field-level determinants of institutional complexity—the presence of various competing logics of decision making—in neoinstitutional theory (Greenwood et al. 2011). The debate builds on Meyer and Scott's (1983) suggestion that organizations in mature, rationalized, and centralized fields face less complexity than organizations in emerging and dispersed fields. To be explicit, due to the persistence and self-reproduction of institutional rules, these properties are not an outcome of the actions of organizations subject to evaluation themselves—which would make the following arguments somewhat circular. Instead, key to understanding evaluative plurality are structural characteristics of organizations' external institutional environment and the actors that materialize and enforce legitimacy in its many facets (DiMaggio and Powell 1983; Scott and Meyer 1994; Zucker 1977).

This literature inspires speculation about the structural determinants of plurality in evaluative landscapes, even though the evaluative practices present do not perfectly reflect a field's institutional environment. Not all the conceivable ways for an organization to be considered successful by its audiences—such as a corporation being judged the most family-friendly place to work or as having the most faithful employees—are necessarily operationalized in some consequential evaluative measure. Notwithstanding, there exist some institutional logics—family and religion—that would back up these types of evaluations (Thornton et al. 2012). Similarly, social and environmental concerns were being used to characterize the corporate sector long before the existence of contemporary corporate social responsibility rankings (Brammer, Jackson, and Matten 2012).

What, then, determines whether the complexity of logics and values actually becomes translated into evaluative plurality? One helpful insight from the sociology of valuation is that the practice of organizational evaluation, as well as the people practicing it, is essential for understanding how evaluative practices are generated, approved, and then enacted in an organizational field. Important questions to ask include the following: Who can legitimately evaluate performance? Who has a stake in and resources for championing or contesting values? And who is the audience of the produced evaluations? As a starting point for future empirical analyses, I suggest to look at cases that can explain who involved in realizing competing values in an organizational field, including the diversity of these organizations' external audiences, the barriers to legitimate participation in the determination of the value of an organization, and the degree to which that value is defined by formal rules and regulation.

As in the preceding discussion, the following propositions are explorative rather than conclusive. In this section, I suggest some of the structural preconditions that foster evaluative multiplicity—such as political contestation and underdeveloped formal regulation. The role of evaluative entrepreneurs in endorsing, delegitimizing, celebrating, and categorizing value is a complementary matter that is, without doubt, worthy of empirical inquiry. The issue has received some noteworthy attention from valuation scholars (Kjellberg et al. 2013; Kornberger et al. 2015; Miller and Power 2013; Pollock and Williams 2016). To what degree evaluative entrepreneurs can contest and construct value, however, is dependent on not only their resources and interests but also properties of the institutional context of organizing, including the values and beliefs of stakeholders, consumers, media observers, state agencies, and professional associations.

One clear determinant of the voices that are involved in the evaluative process is the diversity of the organizational audiences. Organizations with more segmented audiences, that is, those that operate in a variety of industries, regions, or social domains, face a greater diversity of evaluative criteria. Padgett and Powell (2012) offer a compelling example within the biotech industry, which was introduced above. The "dedicated biotechnology firm" applies for patents in its commercial realm, reports profits to the financial realm, and publishes papers in the scientific realm. In alignment with my argument, the authors conclude that organizational diversity "can generate divergent standards and multiple kinds of rules, resulting in competing criteria and gauging success" (Padgett and Powell 2012:438).

Proposition 2a: The more external audiences constitute the institutional environment of a set of organizations, the more plural is the organizations' evaluative landscape.

The presence of various audiences is not immediately analogous with a variety of evaluators. In fact, not every audience can, or wants to, make claims about which behaviors are valued. The breadth of participating evaluators who can make a legitimate claim about evaluation may depend on how taken-for-granted is the concept underlying an organization's performance, as suggested in previous work in both neoinstitutionalism (Greenwood et al. 2011) and the sociology of valuation (Espeland and Stevens 1998). When value is highly institutionalized in an organizational field, experts (rather than users or consumers) pass judgment about quality (Jepperson 2002). Highly mature fields, where value is accepted as fact rather than a matter of taste, can generally be expected to produce evaluative landscapes with more consistent criteria.

Due to a long history of expert theorization, for example, it is clearer what to expect from haute cuisine than from a good food truck (Rao et al. 2003). The guidelines for being awarded a Michelin star are relatively clear, whereas the proper actions required to excel on Yelp or TripAdvisor are uncertain (Orlikowski and Scott 2013). Similarly, until recently, U.S. winemakers were evaluated based on rules that are more standardized than those for beer brewers (Shapin 2012; Zhao and Zhou 2011). In the cases of restaurants and alcoholic beverages, the institutionalization of value is the outcome of social and political processes. In others, the nature of the produced goods and resources in a field can determine how institutionalized is the value. For instance, it is easier to assess the objective quality of semiconductors than the quality of cultural goods, such as books or movies (Karpik 2010; Podolny 2001). When value is less institutionalized, and therefore more open to the interpretations and tastes of users, participation in the evaluation of products and performance is more accessible, and more voices can make legitimate claims about how to evaluate value.

Proposition 2b: The less institutionalized and the more contested are values in an institutional environment of a set of organizations, the more plural is their evaluative landscape.

Following this argument, the form and content of evaluation depends on whether the experiences of users and customers are a valid criterion in the evaluation process or whether decisions are limited to a select few reviewers deemed to have expert knowledge. There may be other "entry barriers" to the evaluation as well. Aspects other than the institutionalization of value and audience diversity can determine who has access to the evaluation process. One important aspect that can restrict access to evaluation is how the specific field or industry is governed (DiMaggio and Powell 1983; Fligstein 1993, 1996; Meyer and Scott 1983). Corporations, to further elaborate on an earlier example, face a more rationalized form of value—that of the shareholder—than do nonprofit organizations, which are often evaluated on a broad range of values and various emerging measures of social impact. Also, in contrast to social impact measures, shareholder value is determined by a rigid and negotiated order of accounting regulations (Barman 2007; Macintosh and Quattrone 2010; Powell et al. 2017).

Formal restrictions on evaluation do not necessarily take the shape of legal definitions—such as laws that determine what profit is and how a firm must report it. They are also frequently defined by professional associations through codes of conduct and formal entry barriers (Abbott 1988; DiMaggio and Powell 1983). According to Abbott (1988), jurisdictional struggles between professions often lead to the creation of authorities charged with overseeing the assessment of quality. Broadly speaking, when rationalized rules regulate value, there are fewer valid ways to evaluate performance (Meyer and Scott 1983). The American Bar Association has established hard rules to define what constitutes the appropriate training and behavior of a professional lawyer, whereas there is no central authority that similarly codifies those of a manager. Consequently, it is a clearer matter to evaluate the success of an aspiring lawyer than that of a junior manager. Despite the differing specific regulations of state bar associations, one might speculate that this is why rankings for business schools are more fragmented than those for law schools.

Proposition 2c: The less rationalized the rules and regulations in the institutional environment of a set of organizations, the more plural is their evaluative landscape.

In consequence, fields with a high diversity of empowered evaluators, little institutionalized value, and few rationalized rules should face a more fragmented evaluative landscape than those with a single audience, high institutional maturity, and high levels of rationalization. This helps explain why law schools are exposed to a more uniform evaluative regime than business schools, why the evaluation of haute cuisine restaurants is relatively standardized, and why nonprofit organizations are evaluated in more diverse ways than for-profit businesses.

Some caveats apply to these explorative propositions about the sources of fragmentation in evaluative landscapes. First, open questions remain about possible interactions between the various processes that generate plurality. For instance, I have argued that rationalization limits the range of applicable modes of evaluation. This is true when the process of rationalization spans multiple audiences, such as in the case of laws that are applicable and known to all audiences. In circumstances where each audience is itself highly rationalized, for example, in its professional methods and standards (McPherson and Sauder 2013), but where there is no overarching process of rationalization that bundles various values into one, evaluative plurality can become even higher (for a related discussion, see Greenwood et al. 2011).

Second, additional processes that enable a larger number of people to legitimately evaluate organizations may be at work. For instance, technological innovation can be a source of empowerment for groups that were previously excluded from the evaluation process. Technologies such as the Internet can expand the population of potential evaluators; the accelerating use of data-aggregation techniques and technical access to shared platforms that allow for the passing of judgment have increased markedly over time (Adkins and Lury 2011; Orlikowski and Scott 2013). Returning to the observation that evaluative entrepreneurs are constrained not only by their own resources—the cost of a review—but also by the taken-for-granted beliefs among organizational leaders and their constituents, it is worth investigating when attempts at rating or ranking fail in spite of such technological affordances.

Finally, future contributions to the sociology of valuation should consider the competition between various evaluators for resources and audience attention. On one hand, whether all conceivable and legitimate evaluation methods would actually be applied by "evaluative entrepreneurs" remains an open question. On the other interorganizational dynamics may be present within evaluative landscapes that put evaluators out of business, make these evaluators change or adapt their criteria, or lead to the creation of new evaluative practices (Kornberger et al. 2015). This question extends beyond the narrow domain of evaluations and to the interchanges between different forms of institutional intermediaries, such as professional norms, government regulation, and social movement pressures (Timmermans and Epstein 2010; Walker and Rea 2014).

Scope Conditions

As components of the institutional environment of organizations, evaluations are a powerful form of soft regulation (Miller and Power 2013; Timmermans and Epstein 2010). Yet, the propositions offered in this article hinge on several crucial assumptions. As for most conceptual work, the scope of the propositions is limited. First, my theory specifically addresses quantitative forms of evaluation and therefore presumes that performance metrics are at least somewhat salient in the organizational field. Some arenas in which qualitative review (such as peer review) prevails may operate differently. These areas that are sheltered from or resistant to quantitative evaluation—such as family or a person's character—are explicitly excluded. Second, this theory presupposes that the subject of evaluation can meaningfully respond to evaluation. This limits the theory to entities that are capable of somewhat independently choosing their structures and practices, a choice that may be compromised in certain highly regulated industries and in the government sector (Carpenter and Krause 2012). This may be true for certain goods—such as films, whose sets are sometimes portrayed as temporary organizations (Bechky 2006; Rossman and Schilke 2014)—but excludes most product ratings. Third, the causal channels of rankings assume a social world in which organizations experience some level of economic or political accountability upon which their survival depends. This likely means that consumers can freely choose where they procure goods and services, and that producers can determine the quantities and qualities of their services and goods.

CONCLUSION

Synopsis

The salience of evaluation is an astonishing characteristic of contemporary life. Assessing bottom lines, appraising reputation, and measuring impact are authoritative practices vis-à-vis consumers, employees, managers, and funding bodies (Colyvas and Powell 2009; Espeland and Stevens 1998; Power 1997). While counting and comparing performances has long been practiced—from business profits to population statistics, university graduation rates, and police arrests—the current public availability and magnitude of performance metrics is unprecedented

(Carruthers and Espeland 1991; Espeland and Sauder 2007; Porter 1996). An emerging sociological literature has established that evaluative landscapes, that is, the universe of rankings, ratings, and standards in an organizational field, can be drivers of the convergence of organizational behavior. Clear-cut cases, such as those of investment banks (Podolny 1993), securities markets (Zuckerman 1999), and law schools (Espeland and Sauder 2007), have been pivotal in furthering the discipline's understanding of how evaluations operate.

Research prompted by these insightful studies has focused on the mechanisms that explain why single evaluative practices affect organizations. The primary contribution of my article is to highlight that these approaches have an important limitation. The observable plurality of evaluative criteria in organizational fields challenges the underlying assumption that fields have uniform social hierarchies rather than competing social orders based on different notions of value. Research shows that the three causal mechanisms commonly discussed as engines of the evaluation effects on organizational fields—reflexivity, categorical legitimation, and status effects—are muffled if external critics disagree about the relevant criteria of evaluation. This observation implies that recognizing the presence of constellations of evaluative practices in an organizational field is crucial to understanding when and why evaluative landscapes shape organizational behavior. A multiplicity of evaluative criteria inhibits evaluative certainty and reduces status anxiety and categorical salience. As such, plurality can be expected to lessen the established effects of external evaluations on organizational behavior.

The plurality of evaluative landscapes determines whether or not evaluative practices will homogenize behavior. To put this problem into context, I suggested that insights be drawn from the sociology of valuation and neoinstitutional theory with regard to competing value systems (Boltanski and Thévenot 2006; Greenwood 2011; Lamont 2012). The explanation for inconsistent evaluation is related not to characteristics of the ranked organizations or specific evaluators, but primarily to the institutional environment of the ranked organizations. The key to understanding both the effects of plurality and its origins is the exploration of the processes for developing evaluations and the properties of the fields in which value is created (Kjellberg et al. 2013; Miller 2001). By comparing a wide range of organizational forms, I propose that the diversity of people who are empowered to pass judgment, the institutionalization of value, and the presence of rationalized rules about how to assess value are preconditions of evaluative certainty. At the same time, organizational outcomes are highly variable in the absence of consistent evaluation criteria. Deviant organizational responses to evaluation range from total disregard to political defiance to strategic differentiation. These responses are plentiful, but they exist in the shadow of current research on organizational evaluation (Greenwood et al. 2011; Sauder and Espeland 2009).

Research Implications

The argument presented here consolidates theories from two relatively distant literatures—organizational theory and the sociology of valuation—and research implications arise for both.

The theory of plurality suggests that there are varying degrees of consistency in the evaluative landscapes of different organizational fields. To date, the body of empirical work has relied heavily on case studies involving highly consistent and salient evaluative landscapes. To better understand the effects of variation in evaluative landscapes, researchers should turn their attention to fields in which value is less institutionalized and rules for evaluation are still evolving, such as emerging markets. One potential research site is the responses of public administrations to the numerous rankings of places (whether cities, states, or countries), which inconsistently measure differentiated values, such as competitiveness, quality of living, and sustainability. These rankings prompt many convergent and divergent responses, including political activism and the co-opting of the evaluation development process. Future research could contribute by systematically documenting deviant behavior to organizational evaluation and by establishing whether status dynamics, such as middle-status conformity, apply to contexts with inconsistent evaluative standards (Phillips et al. 2013; Zuckerman 1999).

Another possible research question that follows directly from this discussion is how some fields are characterized by having more salient evaluations than others. Most sociological studies examine the presence of consequential evaluative indices, but little attention is given to the shaping of organizational fields in the absence of these indices. In fields such as social movement organizations or the informal economy, ratings or rankings are less pervasive. One promising field of study is the nonprofit sector, where large-scale performance measurement is a very recent development and is often incompatible with the culture of the constituent organizations (Frumkin 2005; Hwang and Powell 2009). In organizational studies, exceptions to the norm of treating evaluators as outside the frame of analysis are still rare (e.g., Christensen and Strandgaard Pedersen 2013; Espeland and Sauder 2016; Orlikowski and Scott 2013). However, valuable contributions could be made by additional studies investigating the evaluators themselves and their motivations.

Finally, sociologists of valuation have embraced Karpik's (2010) argument that evaluative practices can serve as judgment devices to facilitate the decision making of audiences in markets with singular goods. However, the various ways that organizations use these same devices for different purposes have attracted less attention. From the perspective of organizations, evaluations are often a *justificatory device* used to defend a given behavior as appropriate (Suchman 1995). As "people constantly mobilize moral principles and views of the common good to talk about the effects of market process" (Fourcade and Healy 2007:18), evaluations can be used as a tool kit of external evidence for organizational performance (Swidler 1986; Thornton et al. 2012). Studies could explore how organizations use external measures of the noneconomic aspects of success, in the form of rankings, ratings, and accreditations, to frame and justify their market-related actions and strategic differentiation.

For students of organizations, the rise of external evaluations of goods, services, and organizations has meant access to fascinating data. Ratings, rankings, and awards can serve as useful proxies for concepts hard to operationalize—from quality and performance to reputation and status. The resulting studies provide indisputable insights for better understanding the unintended effects of evaluations and the behavior of organizations. We would be misguided, however, in ignoring the insight from sociology of valuations that external evaluations are also the outcomes of social and political processes. Here, I have suggested some considerations as to how valuation processes matter in creating and maintaining order in organizational fields in the face of multiple orders of worth and value.

ACKNOWLEDGMENTS

I would like to thank Andrew Cohen, Gregoire Croidieu, Jared Furuta, Mark Granovetter, Jennifer Hill, Markus Höllerer, Candace Jones, Martin Kornberger, John Meyer, Paolo Parigi, Woody Powell, Mitchell Stevens, Cristobal Young, and Xueguang Zhou for their helpful comments. Earlier drafts of this article were presented at the 2014 conference on Materiality in NIT at the Vienna University of Economics and Business as well as the Economic Sociology workshop and Networks and Organizations workshop at Stanford. I also acknowledge the Europe Center at Stanford's Freeman Spogli Institute for its generous travel support.

NOTES

1. A related discussion concerns the performativity of economic theory in science and technology studies, which postulates an "*homo economicus* 'made flesh' by economic technologies" (Fourcade and

Healy 2007:302; also Callon 1998). MacKenzie (2006) proposes three levels of performativity, from generic to effective to Barnesian performativity. While generic performativity "might just be window dressing for processes that might have happened anyway" (Fourcade and Healy 2007), Barnesian performativity indicates that economic actors fully embrace the theory to "better correspond to the model" (MacKenzie 2006). The critical accounting literature in particular has highlighted the performativity of accounting devices—among others, through linking distinct social domains, by adjudicating success and failures, and by establishing a form of soft regulation—in the constitution of economic value (Miller 2001; Miller and Power 2013).

2. Note that prizes only assign prestige for a certain "approved offering" (Goode 1978; Rossman and Schilke 2014) and therefore express only high status. In contrast, rankings assign each evaluated organization to an explicit status position, from high to low. Negative prizes (such as the Golden Raspberry for the worst movie) extend the spectrum of awards into the low-status realm.
3. The diversity of sources of these rankings, ratings, and awards has increased in a similar manner over the 50-year period. For methodological details, see Brandtner and Bromley 2017.

REFERENCES

Abbott, Andrew. 1988. *The System of Professions: an Essay on the Division of Labor*. Chicago: University of Chicago Press.

Adkins, Lisa, and Celia Lury. 2011. "Introduction: Special Measures." *Sociological Review* 59(2):5–23.

Anand, Narasimhan, and Richard A. Peterson. 2000. "When Market Information Constitutes Fields: Sensemaking of Markets in the Commercial Music Industry." *Organization Science* 11(3):270–84.

Baka, Vasiliki. 2015. "Understanding Valuing Devices in Tourism Through 'Place-making.'" *Valuation Studies* 3(2):149–80.

Barman, Emily. 2007. "What Is the Bottom Line for Nonprofit Organizations? A History of Measurement in the British Voluntary Sector." *Voluntas* 18(2):101–15.

Barman, Emily. 2015. "Of Principal and Principle: Value Plurality in the Market of Impact Investing." *Valuation Studies* 3(1):9–44.

Barman, Emily. 2016. *Caring Capitalism*. Cambridge, UK: Cambridge University Press.

Battilana, Julie, and Matthew Lee. 2014. "Advancing Research on Hybrid Organizing: Insights from the Study of Social Enterprises." *Academy of Management Annals* 8(1):397–441.

Bechky, Beth A. 2006. "Gaffers, Gofers, and Grips: Role-based Coordination in Temporary Organizations." *Organization Science* 17(1):3–21.

Beckert, Jens, and Patrik Aspers. 2011. *The Worth of Goods*. Oxford, UK: Oxford University Press.

Berger, Peter L., and Thomas Luckmann. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Anchor Books.

Bermiss, Y. Sekou, Edward Zajac, and Brayden G. King. 2013. "Under Construction: How Commensuration and Management Fashion Affect Corporate Reputation Rankings." *Organization Science* 25(2):591–608.

Blank, G. 2006. *Critics, Ratings, and Society: The Sociology of Reviews*. New York: Rowman & Littlefield.

Boltanski, Luc, and Laurent Thévenot. 2006. "On Justification: Economies of Worth." Princeton, NJ: Princeton University Press.

Brammer, Stephen, Gregory Jackson, and Dirk Matten. 2012. "Corporate Social Responsibility and Institutional Theory: New Perspectives on Private Governance." *Socio-Economic Review* 10(1):3–28.

Brandtner, Christof, and Patricia Bromley. 2017. "Social Evaluations and the Rise of Win-Win Discourse in U.S. Firms, 1960–2010." Unpublished manuscript.

Bromley, Patricia, and John W. Meyer. 2015. *Hyper-organization: Global Organizational Expansion*. London, UK: Oxford University Press.

Bromley, Patricia, and Amanda Sharkey. 2017. "Casting Call: The Social Construction of Firms as Actors, 1960–2010." *Accounting, Organizations and Society* 59:3–20. Retrieved June 30, 2017 (http://www.patriciabromley.com/BromleySharkeyAOS.pdf). Bromley.

Brunsson, Nils. 1989. *The Organization of Hypocrisy: Talk, Decisions and Actions in Organizations*, translated by N. Adler. Chichester, UK: Wiley.

Callon, Michel. 1998. "The Embeddedness of Economic Markets in Economics." Pp. 1–57 in *The Laws of the Markets*, edited by M. Callon. Oxford, UK: Wiley-Blackwell.

Carpenter, Daniel P., and George A. Krause. 2012. "Reputation and Public Administration." *Public Administration Review* 72(1):26–32.

Carruthers, Bruce G., and Wendy N. Espeland. 1991. "Accounting for Rationality: Double-entry Bookkeeping and the Rhetoric of Economic Rationality." *American Journal of Sociology* 97(1):31–69.

Chatterji, Aaron K., and Michael W. Toffel. 2010. "How Firms Respond to Being Rated." *Strategic Management Journal* 31(9):917–45.

Chen, Ya-Ru, Randall S. Peterson, Damon J. Phillips, Joel M. Podolny, and Cecilia L. Ridgeway. 2012. "Introduction to the Special Issue: Bringing Status to the Table. Attaining, Maintaining, and Experiencing Status in Organizations and Markets." *Organization Science* 23(2):299–307.

Christensen, Bo T., and Jesper Strandgaard Pedersen. 2013. "Restaurant Rankings in the Culinary Field." Pp. 235–59 in *Exploring Creativity*, edited by B. Moeran and B. T. Christensen. Cambridge, UK: Cambridge University Press.

Colyvas, Jeanette A., and Walter W. Powell. 2009. "Measures, Metrics, and Myopia: The Challenges and Ramifications of Sustaining Academic Entrepreneurship." Pp. 79–111 in *Measuring the Social Value of Innovation: A Link in the University Technology Transfer and Entrepreneurship Equation*, edited by G. D. Libecap. Bingley, UK: Emerald Group.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48(2):147–60.

Dunn, Mary B., and Candace Jones. 2010. "Institutional Logics and Pluralism: The Contestation of Care and Science Logics in Medical Education, 1967–2005." *Administrative Science Quarterly* 55(1):114–49.

Durand, Rudolphe, and Lionel Paolella. 2013. "Category Stretching: Reorienting Research on Categories in Strategy, Entrepreneurship, and Organization Theory." *Journal of Management Studies* 50(6):1100–23.

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(3):442–76.

Espeland, Wendy N., and Michael Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology* 113(1):1–40.

Espeland, Wendy N., and Michael Sauder. 2016. *Engines of Anxiety: Academic Rankings, Reputation, and Accountability*. New York: Russell Sage Foundation.

Espeland, Wendy N., and Mitchell L. Stevens. 1998. "Commensuration as a Social Process." *Annual Review of Sociology* 24:314–43.

Espeland, Wendy N., and Stacy E. Lom. 2016. "Noticing Numbers: How Quantifications Changes What We See and What We Don't." Pp. 18–34 in *Making Things Valuable*, edited by M. Kornberger, L. Justesen, A. K. Madsen, and J. Mouritsen. Oxford, UK: Oxford University Press.

Fligstein, Neil. 1993. *The Transformation of Corporate Control*. Cambridge, MA: Harvard University Press.

Fligstein, Neil. 1996. "Markets as Politics: A Political-cultural Approach to Market Institutions." *American Sociological Review* 61(4):656–73.

Fligstein, Neil, and Doug McAdam. 2012. *A Theory of Fields*. Oxford, UK: Oxford University Press.

Fourcade, Marion, and Kieran Healy. 2007. "Moral Views of Market Society." *Annual Review of Sociology* 33(1):285.

Frumkin, Peter. 2005. *On Being Nonprofit*. Cambridge, MA: Harvard University Press.

Gardner, William L., and Mark J. Martinko. 1988. "Impression Management in Organizations." *Journal of Management* 14(2):321–38.

Glynn, Mary Ann, Pamela S. Barr, and M. Tina Dacin. 2000. "Pluralism and the Problem of Variety." *Academy of Management Review* 25(4):726–34.

Goldberg, Amir, Michael T. Hannan, and Balázs Kovács. 2016. "What Does It Mean to Span Cultural Boundaries? Variety and Atypicality in Cultural Consumption." *American Sociological Review* 81(2):215–41.

Goode, William J. 1978. *The Celebration of Heroes: Prestige as a Social Control System*. Berkeley: University of California Press.

Gould, Roger V. 2002. "The Origins of Status Hierarchies: A Formal Theory and Empirical Test." *American Journal of Sociology* 107(5):1143–78.

Greenwood, Royston, Christine Oliver, Roy Suddaby, and Kerstin Sahlin-Andersson. 2008. *The SAGE Handbook of Organizational Institutionalism*. London: Sage.

Greenwood, Royston, Mia Raynard, Farah Kodeih, Evelyn R. Micelotta, and Michael Lounsbury. 2011. "Institutional Complexity and Organizational Responses." *Academy of Management Annals* 5(1):317–71.

Gülker, Silke, Dagmar Simon, and Marc Torka. 2012. "Evaluation of Science as Consultancy?" *Quaderni* (1):41–54.

Hannan, Michael T., and J. Freeman. 1989. *Organization Ecology*. Cambridge, MA: Harvard University Press.

Hannan, Michael T., Laszlo Pólos, and Glenn R. Carroll. 2007. *Logics of Organization Theory*. Princeton, NJ: Princeton University Press.

Heimer, Carol A. 2001. "Cases and Biographies: An Essay on Routinization and the Nature of Comparison." *Annual Review of Sociology* 27:47–76.

Hsu, Greta, Michael T. Hannan, and Özgecan Koçak. 2009. "Multiple Category Memberships in Markets: An Integrated Theory and Two Empirical Tests." *American Sociological Review* 74(1):150–69.

Hwang, Hokyu, and Walter W. Powell. 2009. "The Rationalization of Charity: The Influences of Professionalism in the Nonprofit Sector." *Administrative Science Quarterly* 54(2):268–98.

Jepperson, Roland L. 2002. "The Development and Application of Sociological Neoinstitutionalism." Pp. 229–66 in *New Directions in Contemporary Sociological Theory*, edited by J. Berger and M. J. Zelditch. Lanham, MA: Rowman & Littlefield.

Kalberg, Stephen. 1980. "Max Weber's Types of Rationality: Cornerstones for the Analysis of Rationalization Processes in History." *American Journal of Sociology* 85(5):1145–79.

Karpik, Lucien. 2010. *Valuing the Unique*. Princeton, NJ: Princeton University Press.

Kjellberg, Hans, Alexandre Mallard, Diane-Laure Arjaliès, Patrik Aspers, Stefan Beljean, Alexandra Bidet, Alberto Corsin, Emmanuel Didier, Marion Fourcade, Susi Geiger, Klaus Hoeyer, Michèle Lamont, Donald MacKenzie, Bill Maurer, Jan Mouritsen, Ebba Sjögren, Kjell Tryggestad, François Vatin, and Steve Woolgar. 2013. "Valuation Studies? Our Collective Two Cents." *Valuation Studies* 1(1):11–30.

Kornberger, Martin, and Chris Carter. 2010. "Manufacturing Competition: How Accounting Practices Shape Strategy Making in Cities." *Accounting, Auditing & Accountability Journal* 23(3):325–49.

Kornberger, Martin, Lise Justesen, Anders Koed Madsen, and Jan Mouritsen. 2015. *Making Things Valuable*. Oxford, UK: Oxford University Press.

Kovács, Balázs, and Michael T. Hannan. 2010. "The Consequences of Category Spanning Depend on Contrast." Pp. 175–201 in *Categories in Markets: Origins and Evolution*, edited by G. Hsu, G. Negro, and Ö. Koçak. Bingley, UK: Emerald Group.

Kovács, Balázs, and Amanda Sharkey. 2014. "The Paradox of Publicity: How Awards Can Negatively Impact the Evaluation of Quality." *Administrative Science Quarterly* 59(1):1–33.

Lamont, Michelle. 2012. "Toward a Comparative Sociology of Valuation and Evaluation." *Annual Review of Sociology* 38:201–21.

Lapsley, Irvine, Peter Miller, and Fabrizio Panozzo. 2010. "Accounting for the City." *Accounting, Auditing & Accountability Journal* 23(3):305–24.

Luca, Michael. 2011. "Reviews, Reputation, and Revenue: The Case of Yelp.com." Working Paper No. 12-016, Harvard Business School, Boston, MA.

Macintosh, Norman B., and Paolo Quattrone. 2010. *Management Accounting and Control Systems: An Organizational and Sociological Approach*. New York: Wiley.

MacKenzie, Donald. 2006. *An Engine, Not a Camera*. Cambridge, MA: MIT Press.

Mair, Johanna, Judith Mayer, and Eva Lutz. 2015. "Navigating Institutional Plurality: Organizational Governance in Hybrid Organizations." *Organization Studies* 36(6):713–39.

McPherson, Chad M., and Michael Sauder. 2013. "Logics in Action: Managing Institutional Complexity in a Drug Court." *Administrative Science Quarterly* 58(2):165–96.

Meyer, John W., and Patricia Bromley. 2013. "The Worldwide Expansion of 'Organization.'" *Sociological Theory* 31(4):366–89.

Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83(2):340–63.

Meyer, John W., and W. Richard Scott. 1983. *Organizational Environments: Ritual and Rationality*. Beverly Hills, CA: Sage.

Miller, Peter. 2001. "Governing by Numbers: Why Calculative Practices Matter." *Social Research* 68(2):379–96.

Miller, Peter, and Michael Power. 2013. "Accounting, Organizing, and Economizing: Connecting Accounting Research and Organization Theory." *Academy of Management Annals* 7(1):557–605.

Orlikowski, Wanda J., and Susan V. Scott. 2013. "What Happens When Evaluation Goes Online? Exploring Apparatuses of Valuation in the Travel Sector." *Organization Science* 25(3):868–91.

Pache, Anne-Claire, and Filipe Santos. 2013. "Inside the Hybrid Organization: Selective Coupling as a Response to Competing Institutional Logics." *Academy of Management Journal* 56(4):972–1001.

Padgett, John F., and Paul D. McLean. 2006. "Organizational Invention and Elite Transformation: the Birth of Partnership Systems in Renaissance Florence." *American Journal of Sociology* 111(5):1463–1568.

Padgett, John F., and Walter W. Powell. 2012. *The Emergence of Organizations and Markets*. Princeton, NJ: Princeton University Press.

Phillips, Damon J., Catherine J. Turco, and Ezra W. Zuckerman. 2013. "Betrayal as Market Barrier: Identity-based Limits to Diversification Among High-Status Corporate Law Firms." *American Journal of Sociology* 118(4):1023–54.

Phillips, Damon J., and Ezra W. Zuckerman. 2001. "Middle-status Conformity: Theoretical Restatement and Empirical Demonstration in Two Markets." *American Journal of Sociology* 107(2):379–429.

Piazza, Alessandro, and Fabrizio Castellucci. 2014. "Status in Organization and Management Theory." *Journal of Management* 40(1):287–315.

Podolny, Joel M. 1993. "A Status-based Model of Market Competition." *American Journal of Sociology* 98(4):829–72.

Podolny, Joel M. 2001. "Networks as the Pipes and Prisms of the Market." *American Journal of Sociology* 107(1):33–60.

Podolny, Joel M. 2005. *Status Signals*. Princeton, NJ: Princeton University Press.

Pollock, Neil, and Luciana D'Adderio. 2012. "Give Me a Two-by-two Matrix and I Will Create the Market: Rankings, Graphic Visualisations and Sociomateriality." *Accounting, Organizations and Society* 37(8):565–86.

Pollock, Neil, and Robin Williams. 2016. *How Industry Analysts Shape the Digital Future*. Oxford, UK: Oxford University Press.

Pontikes, Elizabeth. 2012. "Two Sides of the Same Coin: How Ambiguous Classification Affects Multiple Audiences' Evaluations." *Administrative Science Quarterly* 57(1):1–38.

Porter, Theodore M. 1996. *Trust in Numbers*. Princeton, NJ: Princeton University Press.

Powell, Walter W., and Paul J. DiMaggio. 1991. *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.

Powell, Walter W., Achim Oberg, Valeska Korff, Carrie Oelberger, and Karina Kloos. 2017. "Institutional Analysis in a Digital Era: Mechanisms and Methods to Understand Emerging Fields." Pp. 305–44 in *New Themes in Institutional Analysis Topics and Issues from European Research*, edited by G. Krücken, C. Mazzo, R. Meyer, and P. Walgenbach. Cheltenham, UK: Edward Elgar.

Powell, Walter W., Kelley Packalen, and Kjersten Whittington. 2012. "Organizational and Institutional Genesis: The Emergence of High-tech Clusters in the Life Sciences." Pp. 434–65 in *The Emergence of Organizations and Markets*, edited by W. W. Powell and J. F. Padgett. Princeton, NJ: Princeton University Press.

Power, Michael. 1997. *The Audit Society: Rituals of Verification*. Oxford, UK: Oxford University Press.

Rao, Hayagreeva, Philippe Monin, and Rudolphe Durand. 2003. "Institutional Change in Toque Ville: Nouvelle Cuisine as an Identity Movement in French Gastronomy." *American Journal of Sociology* 108(4):795–843.

Rao, Hayagreeva, Calvin Morrill, and Mayer N. Zald. 2000. "Power Plays: How Social Movements and Collective Action Create New Organizational Forms." *Research in Organizational Behaviour* 22:237–81.

Rogerson, Robert J. 1999. "Quality of Life and City Competitiveness." *Urban Studies* 36(5/6):969–85.

Rossman, Gabriel, and Oliver Schilke. 2014. "Close, but No Cigar: The Bimodal Rewards to Prize-seeking." *American Sociological Review* 79(1):86–108.

Salganik, Matthew J., Peter S. Dodds, and Duncan J. Watts. 2006. "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market." *Science* 311(5762):854–56.

Sauder, Michael. 2006. “Third Parties and Status Position: How the Characteristics of Status Systems Matter.” *Theory and Society* 35(3):299–321.

Sauder, Michael, and Wendy N. Espeland. 2009. “The Discipline of Rankings: Tight Coupling and Organizational Change.” *American Sociological Review* 74(1):63–82.

Sauder, Michael, and Gary A. Fine. 2008. “Arbiters, Entrepreneurs, and the Shaping of Business School Reputations.” *Sociological Forum* 23(4):699–723.

Sauder, Michael, Freda Lynn, and Joel M. Podolny. 2012. “Status: Insights from Organizational Sociology.” *Annual Review of Sociology* 38:267–83.

Scott, W. Richard, and John W. Meyer. 1994. *Institutional Environments and Organizations: Structural Complexity and Individualism*. Thousand Oaks, CA: Sage.

Shapin, Steven. 2012. “The Tastes of Wine: Towards a Cultural History.” *Rivista di Estetica* 51(3):49–94.

Simon, Dagmar, and Andreas Knie. 2013. “Can Evaluation Contribute to the Organizational Development of Academic Institutions? An International Comparison.” *Evaluation* 19(4):402–18.

Stark, David. 2000. “For a Sociology of Worth.” Working Paper Series, Center on Organizational Innovation, Columbia University. Retrieved June 30, 2017 (http://www.coi.columbia.edu/.stage/pdf/stark_fsw.pdf).

Stark, David. 2011. *The Sense of Dissonance*. Princeton, NJ: Princeton University Press.

Suchman, Marc C. 1995. “Managing Legitimacy: Strategic and Institutional Approaches.” *Academy of Management Review* 20(3):571–610.

Swidler, Ann. 1986. “Culture in Action: Symbols and Strategies.” *American Sociological Review* 51(2):273–86.

Thornton, Patricia H., William Ocasio, and Michael Lounsbury. 2012. *The Institutional Logics Perspective*. Oxford, UK: Oxford University Press.

Timmermans, Stefan, and Steven Epstein. 2010. “A World of Standards but Not a Standard World: Toward a Sociology of Standards and Standardization.” *Annual Review of Sociology* 36:69–89.

Walker, Edward T., and Christopher M. Rea. 2014. “The Political Mobilization of Firms and Industries.” *Annual Review of Sociology* 40(1):281–304.

Zhao, Wei, and Xueguang Zhou. 2011. “Status Inconsistency and Product Valuation in the California Wine Market.” *Organization Science* 22(6):1435–48.

Zucker, Lynne G. 1977. “The Role of Institutionalization in Cultural Persistence.” *American Sociological Review* 42(5):726–43.

Zuckerman, Ezra W. 1999. “The Categorical Imperative: Securities Analysts and the Illegitimacy Discount.” *American Journal of Sociology* 104(5):1398–1438.

Zuckerman, Ezra W. 2012. “Construction, Concentration, and (Dis)Continuities in Social Valuations.” *Annual Review of Sociology* 38:223–45.

AUTHOR BIOGRAPHY

Christof Brandtner is a PhD candidate in sociology at Stanford University and affiliated with the Stanford Center on Philanthropy and Civil Society and the Research Institute on Urban Management and Governance at WU Vienna. His research investigates the expansion of rational principles of organizing in domains where social and democratic values loom large. His work explores the meaning and sources of rationalization in the nonprofit sector, the relationship between openness and bureaucracy in city administrations, and the consequences of transparency and accountability on organizational behavior.