Do theories of organizations progress?*

Gerald F. Davis
Ross School of Business
The University of Michigan
701 Tappan St.
Ann Arbor, MI 48109-1234
Telephone: (734) 647-4737
gfdavis@umich.edu

Forthcoming in Organizational Research Methods
6/6/2010

* I thank Bob Vandenberg and three anonymous reviewers for their constructive and engaging feedback on prior versions of this paper and Jeff Edwards for organizing an invigorating discussion.
Do theories of organizations progress?

Abstract. This paper analyzes the prospects for cumulative theory development in organization theory. Organization theory generally takes organizations (business corporations, non-profits, hospitals, universities, and so on) as discrete and meaningful units of analysis. The revolution in information and communication technologies over the past two decades has made comprehensive time-series data on millions of organizations widely available to scholars around the world. Yet it is not obvious that organization theory has become more precise, more general, or more accurate as a result. I argue that this is due to three difficulties that additional data cannot fully resolve: (a) researchers lack experimental control, limiting their ability to draw causal inferences, and are largely inattentive to the standards for valid quasi-experimental design; (b) organizations are more appropriately construed as tools rather than as natural objects susceptible to “laws;” and (c) the regularities underlying organizational dynamics change over time such that empirical generalizations that are true during one period may be false in a different period. I conclude by suggesting improvements to research practice that would enable organization theory to progress toward greater insights.
Organization theory is roughly a half-century old as a distinct field of the social sciences. When March and Simon wrote the foundational text for organization theory in 1958, it was hardly a field at all: “However much organizations occupy the thoughts of practicing executives and administrators, and however many books for these practitioners have been written about them, the theory of organizations occupies an insignificant place in modern social science” (March and Simon, 1958: 1). Surveying prior work, they stated that “The literature leaves one with the impression that after all not a great deal has been said about organizations, but it has been said over and over in a variety of languages” (5). With the contemporaneous creation of two specialized journals for the study of organizations (Administrative Science Quarterly in 1956 and the Academy of Management Journal in 1958), the subsequent decade saw a rapid proliferation of research. James Thompson, ASQ’s founding editor, synthesized the empirical work of the field’s first decade in his masterful 1967 book Organizations in Action. Ten years later, a half-dozen theoretical paradigms had emerged to elaborate on or contest various aspects of Thompson’s synthesis. The Ford and Carter administrations saw a new theory or two created every year in a sort of Cambrian explosion of creativity. Yet the theoretical flowering of organization theory’s first two decades was arguably followed by three decades of muted theoretical progress or even stagnation. Like symphony orchestras that play a repertoire of a dozen baroque and classical composers year in and year out, organizational research can sometimes appear like a living museum of the 1970s.

What is more surprising is that during this same time period statistical methods, computing power, and data have made quantum advances. Previously, data was perhaps the limiting factor in testing and refining theory about organizations. Getting a reasonable sample size to run cross-sectional correlations could take months or even years. Now, anyone with a Web connection can assemble datasets in moments that are far more expansive than even the
most ambitious dissertation written in, say, 1990. The cover charge on advanced statistical methods has also gone down. From structural equations modeling and fixed effects regressions to network analysis, it is possible for a typical MBA student today to analyze data with tools far more advanced than those available to the geekiest SAS programmer 20 years ago. And advances in computing power allow ever more data-intensive investigations on ever more portable devices. Yet our theoretical debates are largely stuck where they were two decades ago (e.g., Pfeffer, 1993).

In this paper I examine the mis-match between the rapid expansion of our tools and the limited progress of our theories. I argue that the slow progress in organization theory is not due to limitations of data or method but to a mis-apprehension of the subject matter. The problem with organization theory is not that it is imprecise or not general enough—it is that the field has applied the wrong standards of progress. Hambrick (2007: 1346) states that the field of management’s compulsive devotion to theory has actually hindered progress, as indicated by what is published in the elite journals: “The requirement that every paper must contribute to theory is not very sensible; it is probably a sign of our academic insecurity; and it is costing us in multiple ways.” I argue here that researchers’ nearly universal and costless access to extensive data on organizations, coupled with a compulsive requirement that publications “contribute to theory,” has resulted in a theoretical stalemate that stymies progress. I conclude by suggesting an alternative approach to theoretical progress in the study of organizations.

**Organization theory and the wealth of data**

Organization and Management Theory is broadly concerned with organizations and organizing, with a particular (but not exclusive) emphasis on organizations as distinct, countable units of analysis. As a research domain, organization theory has been dominated for roughly three decades by a half-dozen paradigms, each with a distinct focus (see Scott and Davis, 2007
for a review). Contingency theory asks “How do organizations structure themselves to protect their technical core from the turbulence of the external environment?” (Thompson, 1967). Transaction cost economics focuses on where organizations place their boundaries—that is, which inputs are bought on the market, and which are made inside the organizations (Williamson, 1975). Resource dependence theory sees exchanges between the organization and its environment creating power and dependence relations, and catalogs the adaptations organization can use to address their dependence (Pfeffer and Salancik, 1978). Agency theory seeks to understand the ways public corporations are structured to respond to the problems of accountability and control created by dispersed stock ownership (Jensen and Meckling, 1976). Population ecology argues that the structural devices that organizations implement to make themselves reliable and accountable thereby make them resistant to change, and thus studies the birth and death rates of relatively inert organizations rather than their strategies of individual adaptation (Hannan and Freeman, 1977). And New Institutional Theory argues that forces for rationalization in society lead organizations to become relatively homogeneous, due in part to the actions of the professions and the state (Meyer and Rowan, 1977; DiMaggio and Powell, 1983).

**Limited debate among organizational theories.** Although these paradigms vie for attention among researchers, they generally do not compete head-to-head to explain the same specific phenomena. There is usually no critical test to distinguish between them and decide which one is right, in part because they rarely seek to explain the same regularities. They are distinguished more by the questions they ask than the answers they propose. Thus, while transaction cost and resource dependence researchers have provided somewhat divergent explanations for mergers, neither makes predictions sufficiently precise to end up being falsified and replaced by the other (see Scott and Davis, 2007: chapter 9 for a review and contrast of these two approaches). In
some cases the theories are explicitly incommensurable. The animating question for transaction
cost analysts is why organizations choose to make a component within its own boundaries rather
than to buy it on the market. But agency theorists argue that organizational boundaries are
illusory. According to Jensen and Meckling (1976: 310-311), “It is important to recognize that
most organizations are simply legal fictions which serve as a nexus for a set of contracting
relationships among individuals…Viewed in this way, it makes little or no sense to try to
distinguish those things that are ‘inside’ the firm (or any other organization) from those things
that are ‘outside’ of it. There is in a very real sense only a multitude of complex relationships
(i.e., contracts) between the legal fiction (the firm) and the owners of labor, material and capital
inputs and the consumers of output.” From this perspective, the “make or buy” question that is
central to transaction cost theory is not only uninteresting—it is not even meaningful
theoretically. The really interesting questions are about who owns how much of the corporation
(a question that rarely occurs to theorists in other traditions).

The debates among organizational theories thus tend to be at a metatheoretical level when
they occur at all. Scholars may argue about what are the interesting questions—where
organizations place their boundaries, or why corporations have the ownership structures they do,
or what kinds of firms make what kinds of acquisitions, or what accounts for industry death
rates—but theories rarely compete to provide more accurate or precise answers to the same
questions. Instead, researchers tend to choose the tool to fit the job: transaction cost theory for
questions of vertical integration, ecology for birth and death rates, and so on. Without head-to-
head competition, there is little Darwinian selection on theories of organizations.

Theories have a relatively high turnover rate in some fields such as psychology, but
paradigms in organization theory appear to be encased in amber. Meehl (1978: 807) stated that
“in soft psychology theories rise and decline, come and go, more as a function of baffled
boredom than anything else; and the enterprise shows a disturbing absence of the cumulative character that is so impressive in disciplines like astronomy, molecular biology, and genetics.”

In contrast, organization theories rarely disappear, and the half-dozen major paradigms from the mid-1970s still appear on prelim reading lists more than three decades later. Consider the long-standing domain statement for the Organization and Management Theory Division of the Academy of Management: “Specific domain involves building and testing theory about organizations, their members and their management, organization-environment relations, and organizing processes. The area has a rich intellectual heritage. Theoretical advances in organization theory have included strategic choice, resource dependence theory, organizational ecology and institutional theory. More recently we have provided a home for critical, feminist, cognitive and post-modernist theorists…” The fact that specific theories were named indicates that there was an approved cannon of which researchers should be mindful. Moreover, the limited prospect for a critical test among these named theories also helps explain how they manage to hang on: not because they are all true, but because none of them appears ultimately falsifiable, and all have enough features to be congenial to a theoretically eclectic approach. Whether their findings are cumulative is another matter, taken up later.

Although there is little evident selection pressure to weed out weak theories, not all have fared equally well in the competition for researcher attention. One indication of this is the kinds of papers submitted to the OMT division for the annual meetings of the Academy of Management. In 2005, authors submitting papers or symposia were asked to classify their submission using three keywords from a list of 39, to better match papers with reviewers. 11 of the possible keywords were theories (including agency theory, contingency theory, ecology, institutional theory, resource dependence, and transaction costs), and 429 papers and symposia were submitted. The most popular theory by far was institutional theory, which 25.4% of
submitters chose as one of their three keywords, making it the single most popular keyword. The next most prevalent theory was ecology, with a 6.7% market share, and then resource dependence, with a 3.9% share (Davis, 2006). Thus, to the extent that there is theory in organization theory, it is predominantly New Institutional Theory. Surprisingly, although authors chose three keywords each, 59.8% did not choose any of the 11 listed theories, which is a lower prevalence than would be expected if keywords were chosen at random. This is likely to reflect the distinction between manuscripts initially submitted to OMT and those that end up as journal publications.

In a recent analysis of articles published in the *Academy of Management Journal* over almost five decades, Colquitt and Zapata-Phelan (2007) find evidence that, broadly speaking, the average level of theory testing and theory development have both increased continuously among papers published since the 1970s, and that the most recent year studied (2007) showed the highest level on record for theory building, while 1975 represented the lowest level. This seems contrary to my argument for theoretical stagnation after the Cambrian explosion of the 1970s. Yet an examination of Colquitt and Zapata-Phelan’s (2007) coding scheme resolves the discrepancy. Articles were considered high in theory testing when they grounded predictions in existing theory. Articles were considered high in theory building when they introduced a new construct, and moderate if they introduced a new mediator or moderator. The most prevalent types of articles—and those that got the most citations—were studies that introduced tweaks on existing theories, not those that built new paradigms. Notably, their list of the eight “major macro theories” included the six 1970s-vintage paradigms we have discussed, along with two theories from strategy (Colquitt and Zapata-Phelan, 2007: Table 5), and their list of new constructs introduced in *AMJ* were overwhelmingly at the micro level (Table 6)—particularly those that had gone on to be influential.
In short, the evidence suggests that a handful of paradigms from the late 1970s still exercise a predominant influence on organization theory. Published research continues to refer back to these paradigms, and to the extent that there has been theoretical progress, it tends to take the form of qualifications or modest modifications within these paradigms. If there is paradigm consensus in organization theory, it is under the big tent of New Institutional Theory.

**ICTs and the explosion of data availability.** In the past twenty years there has been a development that should have been sufficient to create a sea-change in theory about organizations. Information and communication technologies (ICTs)--in particular, more powerful PCs, the development of the World Wide Web, and the flourishing of data vendors such as Thomson—have created a stunning increase in the availability of high quality time-series data on extremely large samples of organizations. The quality, quantity, and accessibility of data on organizations as discrete units is remarkable. Without too much exaggeration, it is now possible to find out just about everything about the finances and management of every public corporation in the US since roughly the end of the Second World War, essentially for free. For those of us who grew up in the mainframe days, when using archival data meant working for weeks on end (often during the night, when CPU access was best) to stitch together datasets in incompatible formats from disparate sources, the contemporary data environment is like entering a Home Depot of cheap and plentiful information. It would seem that the raw materials are readily available for addressing our most pressing theoretical questions. With the limiting factor of data availability essentially rendered moot, organization theory should be in the midst of a theoretical Renaissance. Indeed, Robert Shiller (2003: 81) states that, thanks to ICTs, economics today is “roughly where astronomy was when the telescope was invented or where biology was when the microscope was invented.”
An example familiar to many organizational researchers, and virtually any active researcher in accounting or finance, is the Wharton Research Data Service (WRDS). WRDS combines a superset of data on firms covered by Compustat, CRSP, and several other sources and makes it accessible via a single easy-to-use portal. Firms are identified by common ID and year variables to make it easy to merge a large number of time series together. WRDS is particularly good at addressing some of the vagaries that plagued individual datasets such as Compustat (e.g., the spotty availability of information on industry segments), as it is designed to serve the needs of researchers rather than people in business, who might be less concerned with historical data. It is easy to use and fast: downloading 50 years’ worth of data on thousands of firms rarely takes longer than a minute. Undoubtedly our current undergraduates will soon be downloading data to their cellphones and running sophisticated fixed-effects regressions on every company listed on the New York Stock Exchange in the 20th century.

The possibilities for new discovery and theoretical advance would appear to be enormous. As a test case, I was curious about how the correlations among different measures of size—sales, assets, number of employees, and market capitalization—had changed since 1950. Organizational size is a fairly basic construct in organization theory, and few papers that study organizations as units of analysis fail to “control for size” in their statistical models. But my hunch was that the various measures used interchangeably to control for size were not in fact interchangeable, and that they may have diverged over time as the US shifted from a manufacturing to a service economy. (GM has many employees, huge revenues, and a miniscule market capitalization; Exxon has huge revenues, modest employees, and a vast market capitalization; Google has few employees, modest revenues, and a huge market capitalization; Wal-Mart is super-sized on all three measures.) Using WRDS, I was able to download data for several thousand firms from 1950 to 2005 inclusive (in Stata format) and answer the question of
how different measures of size have co-varied over time in under 10 minutes. When I was in graduate school during the Crustacean Era, answering this question would have taken literally months, and might have made for a fine dissertation. With WRDS it was an idle exercise, equivalent to the kinds of science experiments one does in introductory chemistry. (My findings are shown in Figure 1, for those currently searching for a dissertation topic.)

**Insert Figure 1 about here**

This example makes two points. First, *size* would seem to be a fairly fundamental concept, yet the most commonly used measures of size have diverged substantially over time, indicating a low level of construct validity. The measures are clearly not interchangeable, and indeed there is little reason to believe that they tap the same underlying construct. Something similar can be said of *age*, which—like size—is used almost reflexively as a control variable by many researchers. Consider the difficulty of measuring the age of the two “largest” US banks in 2010 (as measured by consolidated assets), JP Morgan Chase and Bank of America. JP Morgan Chase is the entity resulting from a 19-year process that started when Manufacturers Hanover bought Chemical Bank in 1991 (renaming itself Chemical); Chemical bought Chase Manhattan (renaming itself Chase Manhattan); Chase bought JP Morgan (renaming itself JP Morgan Chase); and the remaining firm bought several additional banks (including Bank One, which had previously purchased First Chicago NBD, as well as Washington Mutual and Bear Stearns). Bank of America is the entity resulting from 20 years of acquisitions by North Carolina National Bank in Charlotte that included major banks in Virginia, Florida, Missouri, Texas, and elsewhere, with the 1998 purchase of San Francisco’s Bank of America a crowning achievement resulting in a final name change (followed by the 2004 acquisition of FleetBoston and the 2008 purchase of Merrill Lynch). The question “How old is JP Morgan Chase, or Bank of America?” is a conundrum, as a moment’s reflection will show. If *size* and *age* defy simple measurement,
then more subtle aspects of organization (such as “legitimacy” or “power”) are surely even more problematic.

A second point is that big science is not necessarily good science. Effortless access to endless data makes fishing expeditions, such as finding out how measures of size diverge over several decades, trivially easy. What is to restrain researchers from simply conducting a grid search of regressions, driven by no more than a whim? It is seductive to be able to answer almost any question that comes to mind (“What happens if you throw in R&D spending?”), but the hazards for good inference are great.

As a comparison, consider the data required for a paper on hostile takeovers of Fortune 500 companies from 1980-1990 (Davis and Stout, 1992). Several measures came from Compustat (e.g., the market-to-book ratio, return on equity, debt, free cash flow, number of employees)—although at the time the study was done Compustat kept “active” and “inactive” companies (i.e., those that got taken over) in separate databases. Data on the dependent variable (the date of the initial “tender offer”) were hand-coded from microfiche copies of the periodic SEC News Digest (stored at a government library) and, in later years, from quarterly editions of Compact Disclosure. Age of the firm was hand-coded from Moody’s Industrial Manual; the average tenure of employees in the firm’s primary industry was calculated from data in the January 1983 Current Population Survey, accessed via a tape purchased from ICPSR; the names of all directors on each company’s board were typed in by hand from Standard & Poor’s Directory, cleaned, and then processed into board interlock data; information on each firm’s institutional ownership was coded from the Spectrum 13F ownership directory (from another university’s library); and data on the largest shareholding block were hand-coded from still another directory at yet another library. The time from the initiation of the project to its final publication: roughly four years. Answering questions such as “What happens if you throw in…”
in this context could take weeks, and without a strong theoretical prior one was unlikely to pursue a mere hunch. Yet replicating this dataset today for companies from 2000-2010 would take perhaps four days and could be done without leaving the office. The effectively universal availability of high quality time-series data on the financial aspects of US corporations helps explain the character of research in financial economics today, where discussions of theory and data are typically brief: if the question is worth answering, finance researchers can head directly to the regressions, and most of the action is in the econometrics and the robustness checks.

It is not just the financial aspects of public corporations that yield such data. Those interested in non-profits can draw on GuideStar.org, which provides data on 1.8 million non-profit organizations in the US, typically including information on the staff, board of directors, mission, sources and uses of funds, and other relevant facts drawn from annual tax filings. Students of social responsibility can use KLD, providing annual information on the social and environmental aspects of major corporations. And the data bonanza is not limited to the US, as comparable information about organizations around the world is increasingly available, along with over a half-century of annual data on the countries themselves via the World Bank’s World Development Indicators and other sources.

Even more data are available at levels of analysis below that of the organization. Those interested in the formal structure of organizations can draw on enterprise resource planning (ERP) data, as many organizations have centralized repositories with real-time information about their operations, transactions, and human resources. Researchers interested in the informal organization have found that email servers can provide data on networks of workflow, information exchange, and social ties far superior to what one might gather through a survey or direct observation. I am currently working with scholars in architecture, sociology, and information systems to study the links between physical structures, social networks, and
innovation in knowledge-intensive settings such as research labs. From digital floor plans to grant proposals to co-authored papers and patents to their subsequent citations, it is possible to gather a wealth of information unobtrusively to gain insights into the dynamics of a knowledge economy. In the contemporary digital world, organizations are made of data the way that Willy Wonka’s factory is made of candy. It is as if every physicist had a supercollider dropped into his or her backyard.

Of course, plentiful and accessible data need not be good data. Emails are a measure of social networks, but not the measure: if people reserve really important communication for face-to-face meetings, or shout across the cubicle to nearby co-workers (who may be collaborators), then using email data to measure networks still requires a measurement theory. (Similarly, IRS data from tax returns may not provide unerring indicators of income.) But compared to survey data with low response rates and various respondent biases, the unobtrusive data thrown off by organizational processes has much to recommend it.

**Do more data lead to better theory?**

It is safe to say that theory in astronomy improved with the invention of the telescope, and that theory in biology was enhanced by the availability of the microscope. Has organization theory improved as a result of this new avalanche of data? Sadly, and surprisingly, no. Instead, there have been two broad responses to our embarrassment of riches.

One response is statistical fetishism. Based on my reviewing and editorial experience with a large sample of elite journals for organization theory over a two-decade time period, I would argue that the entry charge for publication in top journals of organization theory now seems to include time-series data on large samples with lots of control variables using the latest implementation of Stata. The theoretical requirement, however, has not become correspondingly exacting, and with very large samples come very large sets of statistically significant effects—
which, as Meehl (1978: 822) points out, are almost inevitable: “if you have enough cases and your measures are not totally unreliable, the null hypothesis will always be falsified, regardless of the truth of the substantive theory.” In a strikingly perceptive discussion over a decade ago, sociologist Aage Sørensen lamented the “fascination, if not an obsession, with statistical models and concerns, and a neglect of the need to develop sociological models mirroring conceptions of mechanisms of social processes” leading to statistical models with “a conceptually meaningless list of variables preventing any kind of substantive conclusion” (Sørensen 1998: 238-9, 243).

Just as an all-you-can-eat buffet may not promote good dietary practice, plentiful access to data need not promote good research practice. Ironically, one reason for this is the insistence of management journals that published articles make a “contribution to theory.” Don Hambrick (2007) makes the case forcefully that “the gatekeepers for the top journals in management first screen manuscripts for basic readability and technical adequacy, and then they apply one pivotal test, above all others: Where’s the theory?” Yet theory often serves as a thin rationale for including a particular variable in a regression that—when the sample size gets large enough—almost inevitably turns out to be significant in one direction or another. Given the diverse predictions of the many paradigms in organization theory, it is almost always possible to find a theoretical rationale for a result—either before or after the results are known.

This is not to argue that studies of organization should start with pure theory and then find the precisely appropriate context to study it. There is much to be said for problem-driven research intended to describe and understand variation in the world of organizations (Davis and Marquis, 2005). But it is hard not to conclude that the priorities reflected in the published record stem more from the availability of large-sample datasets than from the substantive importance of the research problems. The ratio of published studies on biotechnology (a tiny but modestly important industry) to published studies on Wal-Mart (perhaps the single most important private
organization in America) is surely one hundred to one at this point. Wal-Mart currently has 1.4 million US employees, making it by far the largest private employer in the country, while the entire biotech industry in the US employs perhaps one-tenth this number (see http://www.bls.gov/oes/2007/may/chartbook.htm). But Wal-Mart would yield a case study, while biotech allows time-series regressions on alliances, initial public offerings, patents, and so on.

A second response to the data deluge has brought, if anything, even less theoretical precision, as New Institutional Theory (NIT) has become the dominant approach to studying organizations. As I noted previously, one out of four papers submitted to the OMT division in 2005 claimed to be a contribution to institutional theory, a bigger market share than all other paradigms combined, and more recent years yield similar results. One of NIT’s foundational articles, “The iron cage revisited” (DiMaggio and Powell, 1983), has become the most widely-cited paper ever published in the American Sociological Review, the flagship journal of the American Sociological Association. NIT is, in effect, the default theory of organizations today in the same way that March and Simon’s (1958) approach was the default theory of the early 1960s and contingency theory was the default theory of the late 1960s and 1970s.

One can entertain many hypotheses for why NIT, as represented by “The iron cage revisited,” has become the dominant theory of organizations. The article cites dead German theorists like Weber and Marx early and often (cf. Stinchcombe, 1982); its imagery of relentless pressures for conformity is appealing in a world in which every strip mall in every town is populated with the same retailers, restaurants, and latte parlors; it includes a simple three-part typology; and it is short—a mere 14 pages, including references. But as one might expect from Murray Davis’s (1971) insightful analysis, its influence cannot be attributed either to its theoretical precision or to its accuracy in accounting for observed facts, because many of its core
propositions cannot be tested now or in the foreseeable future. More than half of the paper’s stated hypotheses are about processes in which the organizational field is the unit of analysis. “By organizational field, we mean those organizations that, in the aggregate, constitute a recognized area of institutional life: key suppliers, resource and product consumers, regulatory agencies, and other organizations that produce similar services or products” (DiMaggio and Powell, 1983: 148). The authors go on to helpfully note that “The structure of a field cannot be defined a priori but must be defined on the basis of empirical investigation. Fields only exist to the extent that they are institutionally defined.”

Armed with this definition of the unit of analysis, researchers presumably should be able to follow standard research practice: define a population of organizational fields, create a sampling frame for that population, and select a representative sample of fields. Next comes measurement of the dependent variable. In plain English, the dependent variable for the field-level hypotheses is the extent to which organizations within a field have come to look like each other, that is, the extent to which they are “isomorphic” on some (unspecified) dimensions. The hypotheses are laid out compactly (DiMaggio and Powell, 1983: 155-156):

B1: The greater the extent to which an organizational field is dependent upon a single (or several similar) source of support for vital resources, the higher the level of isomorphism.

B2: The greater the extent to which organizations in a field transact with agencies of the state, the greater the extent of isomorphism in the field as a whole.

B3: The fewer the number of visible alternative organizational models in a field, the faster the rate of isomorphism in that field.

B4: The greater the extent to which technologies are uncertain or goals are ambiguous within a field, the greater the rate of isomorphic change.

B5: The greater the rate of professionalization in a field, the greater the amount of
institutional isomorphic change.

B6: The greater the extent of structuration of a field, the greater the degree of isomorphics.

Note that across the six hypotheses, the dependent variable is never phrased the same way twice. Is a “higher level of isomorphism” within an organizational field (that is, among those organizations constituting a “recognized area of institutional life”) the same as a “greater rate of isomorphic change” or a “greater degree of isomorphics”?

Assuming we can resolve this question, the next step is to measure the relevant variables. If one has selected a large and representative sample of organizational fields, and calculated the degree of isomorphics in each of those fields, then testing Hypothesis 6 entails measuring the extent of structuration in the field. Again, the authors give a helpful definition to guide measurement: “The process of institutional definition, or ‘structuration,’ consists of four parts: an increase in the extent of interaction among organizations in the field; the emergence of sharply defined interorganizational structures of domination and patterns of coalition; an increase in the information load with which organizations in a field must contend; and the development of a mutual awareness among participants in a set of organizations that they are involved in a common enterprise” (DiMaggio and Powell, 1983: 148).

Note that the hypotheses as stated cannot be meaningfully tested using only a single field, such as health care, the chemical industry, higher education, or biotechnology (although these are the sites of some of the most influential studies). To know whether fields with greater structuration experienced a faster rate of isomorphic change than fields with lesser structuration, we need to observe the rate of isomorphic change over time, as well as the degree of structuration, across a meaningful sample of organizational fields. Of course, that only establishes a correlation; according to the theory, we also need to include measures of
dependence on “a single (or several similar) sources of support for vital resources,” transactions with agencies of “the state,” the number of visible alternative organizational models, technological uncertainty and goal ambiguity at the field level, and the level of professionalization, in order to capture the pure effects of structuration on the rate of isomorphics.

After almost three decades during which NIT has come to be the dominant theory of organizations, and DiMaggio and Powell (1983) has become the most widely-cited paper in the field of sociology, I would be hard pressed to point to a single study that has ever tested any of its field-level hypotheses as they are stated in the theory. Moreover, it is almost impossible to visualize what such a test might ever look like, given the vagueness of the unit of analysis, the ambiguity of the dependent variable, and the distance of many of the core constructs from any measurable indicators. For all practical purposes, this theory cannot be tested or corroborated as it is written. One might imagine efforts to falsify aspects of the theory, but in the rare case where it is attempted, theories such as NIT prove to be a moving target because there is so much wiggle room in the theorizing. Kraatz and Zajac (1996) published a brilliant study documenting that US liberal arts colleges—one of the most “institutionalized” and “structurated” fields, whose constituents’ average age was nearly a century and whose commitment to liberal education (as opposed to mere professional or vocational training) was ironclad—were highly responsive to the demands of their student “customers” during the 1970s and 1980s. Most added “illegitimate” professional and vocational degrees to their offerings, in spite of their mandate; they did so in varied ways that made them more dissimilar over time; and those that did so prospered relative to those that did not. The field, in other words, shifted dramatically in a direction that should have been ruled out by institutional pressures. Yet surprisingly enough, institutional theory lives on, in spite of its evident falsification.
**A problem of theory or a problem of organizations?**

In short, while the data available have become extravagant and the statistical methods have become exotic, organization theory has become even less precise over the past three decades. How has this happened? I suggest that the problem may be due less to the theories than to the subject matter. First, a lack of experimental control means that we are left with quasi-experiments at best. There are widely-known quality standards for quasi-experiments that distinguish stronger research designs from weaker ones (Cook and Campbell, 1979), but these standards are not well integrated into organization theory, as relatively few researchers recognize that their studies are, in fact, quasi-experiments. Instead, researchers commonly rely on long lists of under-theorized control variables which, in practice, degrade rather than improve the kinds of inferences that can be made (cf. Wiliamson, Vandenberg, and Edwards, 2009). Second, it is also possible that organizations are not the types of entities amenable to theories with the kind of precision Meehl (1978) had in mind. In this sense, our problem is similar to that of evolutionary biologists. We don’t ask evolutionary biologists to predict “with precision” what species will thrive in the Galapagos Islands in 50 years, when the ambient temperature has increased by 5 degrees. Third, it has become clear that statistical relations among core constructs in organization theory are often unstable over time because they are historically specific. This is a variant of Hume’s famous “problem of induction,” a problem that no amount of data or theory can cure. I discuss each of these in turn.

**The hazards of unrecognized quasi-experiments.** As in many areas of the human sciences, organization theory is not an experimental science. Researchers who want to know whether some kinds of organizational structures are more effective than others, for instance, or whether corporate social responsibility enhances financial performance, are not generally in a position to randomly assign some organizations to the “treatment condition” while others serve as a control
group. (We will ignore the more obvious flaws of corporate case studies that work backwards from superior performance to infer its causes, such as an unnamed examination of how Circuit City and Fannie Mae went from good to great.) Instead, they will typically select a sample of organizations with the structure or practice in question and a comparison group without and test for differences in performance. But there are hazards to this approach that are rarely recognized, and even more rarely addressed.

When data were limited, researchers often relied on cross-sectional comparisons. Such studies were subject to criticism based on the likely endogeneity of the process being examined. A correlation between social responsibility and profit might exist for any number of reasons: perhaps more profitable firms gave more generously than less profitable firms, so financial performance is a cause, not a consequence, of corporate social responsibility (see Margolis and Walsh, 2003). (Indeed, among firms in the Twin Cities that pledge 5% of their profits to charity, correlations would show a nearly perfect connection between social responsibility and profitability!) Problems of endogeneity are one reason why cross-sectional studies have nearly vanished from organization theory. Any competent reviewer can typically field a half-dozen alternative explanations for observed results that authors are unable to rebut without better data.

But time series data alone are not sufficient to address problems of endogeneity. Most studies of organizational performance are quasi-experiments, although they may not be recognized as such by their authors. What are the performance consequences of increasing diversification, restructuring the board, adopting Total Quality Management, firing the CEO and bringing in an outsider, or implementing layoffs? Each of these is a “treatment” in a quasi-experiment. Cook and Campbell (1979) describe a small number of valid designs, and a large number of flawed designs, and catalogue a number of threats to validity associated with each. We have already touched on problems of the reliability and validity of measures such as size and
age and the problems raised by data dredging, which can undermine statistical conclusion validity (Cook and Campbell, 1979). Other common difficulties that are not resolved simply by using time-series data include effects due to history, maturation, and regression to the mean; thus, performance improvements attributed to hiring a new CEO may be due to history (the industry was in a cyclical low point when the CEO was hired), maturation (the firm was in a transition period that predictably reduced performance prior to the CEO’s hiring), or regression (bad luck beyond the control of the firm had temporarily reduced performance).

Lamentably, weak designs are rife in the literature. This is due in large part to the fact that authors and their reviewers rarely recognize when their study is a quasi-experiment, and thus fail to apply the appropriate standards and address the characteristic weaknesses identified by Cook and Campbell (1979) and others. Perhaps the most basic problem is that organizations assign themselves to “treatments,” and the kinds of organizations that assign themselves to a particular treatment are likely to be different in unobservable ways from those that do not. As Scott and Davis (2007: 312) describe the problem:

Does having a multi-divisional (M-form) structure improve the performance of large firms? In the ideal case, researchers would randomly assign a sample of large firms to two conditions, M-form and placebo. Neither the researchers nor employees in the firms would know which condition they were assigned to [i.e., a “double blind” design]. After two years, their performance would be assessed, and the envelope telling which firms had the M-form and which the placebo would be opened. Practically speaking, researchers are likely to end up comparing a set of M-form adopters and non-adopters, and the adopters (like other firms making major changes) may have simultaneously gotten a new CEO, changed compensation systems, made a few acquisitions and divestitures to balance their portfolio, and retained a consulting firm specializing in effective organizational change…

Indeed, as Cook and Campbell (1979) note, treatments may not be equivalent, creating a further threat to validity: we do not know exactly how big a “dose” of M-form, or re-engineering, or TQM different firms received, even when their new mission statement proclaims the firm’s commitment to quality. In such a situation, how does one discern what the “effective ingredient”
in the treatment is?

Scholars commonly attempt to address this problem by including a number of control variables. For instance, it is obvious that CEOs are likely to be fired in the wake of poor performance, so studies of the effect of a change in CEO on performance will “control” for prior performance. Why is this not sufficient? First, like size, performance has many dimensions, and may not reduce to a single factor. The many, many measures of performance in the literature (return on assets, total market returns, growth in revenues, death rates) lack construct validity: they do not obviously measure the same underlying construct. But second, by failing to account for the process of assignment to treatments, the statistical results are almost certain to be misleading. The same factors that influence assignment to treatments also influence the outcome of interest; that is, they are inherently simultaneous equations. As Christopher Achen puts it, in this situation “…regression gives the wrong answer no matter how much data is available” (Achen, 1986: 22). He concludes, “With quasi-experimental data derived from nonrandomized experiments, controlling for additional variables in a regression may worsen the estimate of the treatment effect, even when the additional variables improve the specification” (Achen, 1986: 27).

The rampant and relatively unthinking use of control variables has come under fire in recent years. Control variables are generally regarded as benign, either ruling out alternative interpretations (in the best case) or doing little harm beyond using up degrees of freedom (in the worst). But Edwards (2008: 481-482) argues that “Control variables do not merely render tests of substantive variables more stringent. Rather, they change the meaning of effects ascribed to substantive variables.” When measured with error (as they nearly always are), control variables can bias the results for substantive variables. Moreover, even when control variables are measured without error, the interpretation of substantive variables is altered when control
variables are included. The problem of interpretation increases the greater the measurement error of the control variables and the greater the number of control variables, even if the substantive variables themselves are measured without error (Williams, Vandenberg and Edwards, 2009: 584). As described in a recent review by Williams et al. (2009), it is common practice for researchers to give little sense of the measurement theory behind control variables or their causal relation to the substantive variables—as causes, consequences, mediators, moderators, or purely exogenous factors. “The net result is the readership not being given a firm conceptual understanding as to why a given control variable was included and why its absence would hinder an unambiguous interpretation of the underlying results if its influence was not removed. Further, readers are exposed to operationalizations of the control variables that possess unknown or poor measurement qualities. Thus, the real possibility exists that the measures of the control variables are conceptually invalid and thus are not representing the control variable construct as stated by the researcher” (Williams et al., 2009: 584-585). The problem is greatly exacerbated by the easy availability of variables that vaguely represent constructs alluded to by prior researchers, or by theory, or common sense, or even whim, encouraging the inclusion of impressively long lists of control variables. In other words, more data makes things worse, not better.

Scholars are not completely oblivious to this issue, of course. Within economics a virtual industry has sprung up around the idea of “naturally occurring experiments,” in which exogenous factors “assign” firms or people to different conditions, avoiding the perceived need to include spurious control variables. For example, companies incorporated in different states occasionally end up fortuitously assigned to different “treatments” when their legislatures pass laws. Thus, scholars found that when states passed antitakeover laws, managers of local firms reduced their ownership stakes in the firms compared to managers of firms in states that did not pass such
laws, suggesting that maintaining control was one rationale for greater managerial ownership (Cheng, Nagar, and Rajan, 2005). One of the most ingenious natural experiments occurred in a squatters’ settlement in Buenos Aires. Home ownership is well-known to be correlated with a number of virtues: home owners are more likely than similar renters to vote in elections, know the names of their school board members, maintain their homes, plant flowers, and successfully raise children who graduate from high school. But the kinds of people who buy homes are different in many hard-to-observe ways from non-homeowners, and thus it is difficult to discern the “pure” effects of home ownership. In a neighborhood of Buenos Aires that used to be a landfill, hundreds of squatters built homes on “reclaimed” land, and a few years later, the government was able to secure the titles to some of the land on which the homes were built—but not all. Thus, some squatters were essentially randomly assigned to the “home ownership” treatment, while similar others were not. Several years later, those that gained titles to their homes differed on a number of dimensions from their untitled neighbors (e.g., in how well the homes were maintained and in their children’s’ school attendance), indicating that ownership per se had a significant effect (DiTella et al., 2007).

Naturally-occurring experiments are a powerful means to ferret out causal effects. But the kinds of organizational questions that can be addressed via natural experiments are somewhat limited, and our knowledge of, for instance, the performance effects of CEO firings remains hazy—at least until I can get funding for an experiment I have in mind.

**Organization theory as (un)natural history.** Organization theorists have a long tradition of treating organizations as if they were natural objects. In the foundational text for the discipline, March and Simon (1958: 4) stated that “the high specificity of structure and coordination within organizations…marks off the individual organization as a sociological unit comparable in
significance to the individual organism in biology.” This biological analogy served organization theory well for several decades, with a head office, hired hands, members, and so on. It has also spawned theory that takes the analogy quite literally, drawing from evolutionary biology texts to derive predictions about organizational birth and death rates (Hannan and Freeman, 1977). But one could as easily have described organizations as being like diesel trucks, with the CEO in the driver’s seat, revenue as the fuel, strategy as a route map, share price as the compass, and so on. That is, rather than being natural objects susceptible to science, they could be perceived as tools susceptible to engineering (Davis and Marquis, 2005).

The idea of a science of diesel trucks, complete with axioms and precise theories, seems fanciful. Yet it is not difficult to imagine scholars pursuing something much like this, using measurable variables (engine displacement, cab length, mileage, brand, color, failure rates) in regressions yielding significant results and quasi-theoretical interpretations. A half-dozen paradigms could arise and stake out distinct domains: a theory of mileage to explain which vehicle is most efficient, a theory of reliability to explain death rates, a theory of brands to explain isomorphism.

Meehl (1978) argues that the distinctive nature of personality as a domain of study makes it “more similar to such disciplines as history, archeology (historical), geology, or the reconstruction of a criminal case from police evidence than the derivation of molar gas laws from the kinetic theory of heat or the mechanisms of heredity from molecular biology.” Something similar can be said of organizations, but with the caveat that there are regularities, if not laws, that can be usefully discerned. Geographer David Harvey notes that capitalism inherently favors novelty: “Its developmental trajectory is not in any ordinary sense predictable, precisely because it has always been based on speculation—on new products, new technologies, new spaces and locations, new labor processes…and the like….There are laws of process at
work under capitalism capable of generating a seemingly infinite range of outcomes out of the slightest variation in initial conditions or of human activity and imagination” (Harvey, 1990: 343). The production of innovative social forms in some sense defies the project of deriving precise or predictive theory. One might have predicted the eventual bankruptcy of General Motors in 1990, and it is possible that one could have anticipated that some means would arise to search across computers connected via networks. But it is highly unlikely that one would have predicted the rise of a company such as Google, much less any aspect of its path of expansion. Facebook? Twitter? Flash mobs, or democratic social movements in Iran organized through Facebook and Twitter? Novel technologies interact with novel social forms such as organizations and social movements in unpredictable ways, and while we might make broad predictions (“most people will have access to the Internet within a decade”), precise predictions about the future structure and prevalence of organizational forms are unlikely to be productive.

A more modest ambition for organization theory, then, is an organizational analogue of natural history: making comprehensible the developmental pathways of organizations and organizing ex post. This will be unsatisfying for those whose model of social science is physics, but is less prone to disappointment due to its imprecision. The appropriate standard would not be predicting what new species emerge, but rendering the ecosystem comprehensible. “Precision” is not the relevant criterion in this situation; “insight” is. Hambrick (2007: 1350) suggests that in place of the requirement that a paper make a contribution to theory, it meet this criterion: “Does the paper have a high likelihood of stimulating future research that will substantially alter managerial theory and/or practice?” There is substantial merit in this proposal.

**Hume again.** A third and ultimately more troubling reason for the lack of theoretical progress in organization theory is that the “laws of nature” do not sit still long enough to be documented.
Like a cadaver that keeps jumping up from the autopsy table, the empirical generalizations derived from the study of organizations often get away from us as time moves on.

Consider the so-called “market for corporate control,” that is, the process by which corporations are taken over by outsiders. This is highly consequential to managers and theoretically central to agency theory, and has critical policy implications. Agency theorists argued that an active takeover market was essential to ensure that corporate managers pursued wise strategies because it provided a concrete punishment for having a low share price (e.g., Jensen and Meckling, 1976). In a typical takeover, an outside bidder makes an offer to a corporation’s shareholders to buy their shares at a premium over their current value, with the intention of buying enough shares to be able to exercise control. Acquirers often fire the target’s top management and install their own team. During the 1980s, 28% of the Fortune 500 received takeover bids, most hostile, and most successful, leading to a wholesale change in the composition, strategies, and industrial structure of America’s largest corporations. A question of great practical and theoretical interest, therefore, is, “What makes companies susceptible to takeover?” As it happens, scholars have studied this question on essentially identical samples (the Fortune 500) using essentially similar statistical models in each of three decades, the 1960s, the 1980s, and the 1990s. During each of these periods there was a wave of mergers, each larger than the last. Yet analyses revealed markedly different dynamics each time. During the 1960s, corporations that were takeover targets were distinguished by low stock market valuations and poorly-connected boards of directors (Palmer et al., 1995). The result of this merger wave was the growth of highly diversified conglomerates such as ITT, Litton Industries, LTV, and Gulf & Western. During the 1980s, targets also had lower valuations than non-targets, but now targets were more highly diversified, while the composition of the board made no difference, in part because the financial markets created innovations such as junk bonds that made it easier to
overcome resistance (Davis and Stout, 1992). In short, the acquirers of the 1960s became the targets of the 1980s, resulting in the dismantling of conglomerate firms and a generalized focus on firms’ “core competence” in the service of shareholder value. During the 1990s, in contrast, the main factor distinguishing targets was that they were in industries that were consolidating (e.g., defense, energy, banking); neither market valuation, level of diversification, nor the composition of the board of directors made a difference. The unsatisfying answer to the question “Why do firms get taken over?” is “It depends (but not in a predictable way).”

The difficulty is an instance of David Hume’s “problem of induction.” As Hume described it in 1748’s *An Enquiry Concerning Human Understanding*, “all inferences from experience suppose, as their foundation, that the future will resemble the past, and that similar powers will be conjoined with similar sensible qualities. If there be any suspicion that the course of nature may change, and that the past may be no rule for the future, all experience becomes useless, and can give rise to no inference or conclusion.” In other words, no matter how systematic and precise our knowledge of the past, we have no logical basis to assume that our inferences will hold true in the future because we cannot guarantee that the laws of nature will not change. To put it in the words of investment prospectuses everywhere, “Past performance is no guarantee of future results.”

Hume’s point was not that we should abandon the lessons of experience and live in a state of perpetual paralysis, not knowing whether the law of gravity will stop working at any moment. It was that we cannot logically assume that today’s laws, based on yesterday’s experiences, will still hold good tomorrow. In the natural world, breaks in the order of things are rare. In the world of business organizations, however, they are common.

The problem of induction is not merely an amusement for philosophers. According to some, it was a major contributor to the US mortgage crisis that precipitated the global economic
downturn that began in 2008. Most mortgages in the US are “securitized,” that is, sold by their originators (the bank or mortgage company that makes the initial loan to the home buyer) to Wall Street banks that bundle them together with thousands of other mortgages and slice them into bonds. These mortgage-backed bonds are divided into tranches based on their risk, which is largely determined by the estimated likelihood that the homeowner will stop making payments (default). When homeowners default, their home typically enters foreclosure, which is costly for the lender and harmful for the home buyer. For essentially the entire history of the American mortgage market prior to 2006, the major predictors of default were that the homeowner did not have sufficient income to pay all of their bills and they had reached the limit of their other sources of credit. Default, in short, was the last resort after all the credit cards were maxed out and no other sources of debt financing were available.

Beginning in 2007, however, the old statistical models predicting default stopped working, and a new set of factors became the significant predictors. Homeowners had now stopped paying their mortgages because the sum of their first and second mortgage debt exceeded the imputed market value of the house—that is, they owed more than the house was worth. Notably, many of those in default had sufficient income to make their mortgage payments, but elected not to. In the past, solvent homeowners almost always continued to make mortgage payments, even when they had negative equity, because of concerns about their credit rating and the social stigma of default. But the CEO of Bank of America stated, “There’s been a change in social attitudes toward default...We’re seeing people who are current on their credit cards but are defaulting on their mortgages.” A debt collector in Mumbai echoed this: “People are walking away from their homes and hanging onto their credit cards, because this is their lifeline.” As a result of this nationwide change in attitude, bonds backed by mortgages that had been rated AAA turned out to be much riskier, prompting widespread downgrades by the bond-
rating agencies. A managing director at Moody’s stated that relying on their pre-2007 models to predict post-2007 behavior was “like observing 100 years of weather in Antarctica to forecast the weather in Hawaii.” The near-universal refrain on Wall Street was that nobody saw this coming; certainly, the number and size of failed or badly distressed financial institutions (AIG, Merrill Lynch, Bear Stearns, Lehman Brothers, Wachovia, Washington Mutual, Fannie Mae, Freddie Mac, Countrywide, and many more) suggests that many smart people with money at stake were as surprised by the turn in the mortgage markets as they would have been if the sun began rising in the west. We are all familiar with the consequences for the global economy.

Expecting to observe empirical generalizations which are consistent over time may be a vain hope for those aspects of organizations beyond the most trivial. Social and technological changes can alter the context that organizations operate in as radically as the asteroid that hit the Earth and changed the developmental paths of species 65 million years ago. What’s the optimal span of control? Six…unless you have an enterprise resource planning system, in which case it may be hundreds. When should a company make an input rather than buy it? When asset specificity is high…or the Web enables detailed transmission of specifications to potential vendors around the world. Making our theory of diesel trucks more precise is of limited value in a world where trucks are powered by electric motors, or tiny nuclear plants under the hood.

This is not meant to justify intellectual nihilism or the abandonment of hope for any social science. There is still value in documenting regularities in the social world, and some hold up quite well. But we should have humility about how long-lasting our accounts might be, particularly when our object is social structures such as organizations.

**Conclusion: Progress in organization theory?**

The upshot of this analysis is that seeking increasingly “precise” or “general” theories about organizations is probably a pointless endeavor. We are drowning in a sea of data, much of
it conveniently formatted, which eliminates a traditional constraint on building knowledge about organizations. But increased access to data has encouraged bad research practice: the use of indicators linked to constructs by weak or absent measurement theory; the rampant use of poorly-rationalized control variables that harm the prospects for inference; and, almost certainly, widespread data fishing. Journals encourage submissions to make a “contribution to theory,” but such contributions almost never take the form of falsification. Instead, a half-dozen paradigms maintain hegemony year after year, facing little danger that new evidence will pile up against them, with New Institutional Theory at the head of the class.

Three central problems stand out. First, researchers typically lack experimental control because for-profit firms generally cannot be assigned to “treatment” and “control” conditions. Studies often take the form of quasi-experiments, but researchers and reviewers rarely recognize that this is the case, leading them to neglect the standards for valid inference described by Cook and Campbell (1979) and their followers. Second, given that organizations are human constructs, like tools, and often operate in a context where innovation and novelty are prized, the prospect of being predictive is limited. The past and present can be made understandable with reference to general processes and mechanisms, but the future is largely off-limits. Finally, organization theory, along with many of the other human sciences, is profoundly limited by the problem of induction described by Hume. Empirical generalizations often change from era to era, or even year to year, so that theories based on past experience are inevitably susceptible to obsolescence. The fact that global investors with real money at stake suffer from this limitation suggests that the social sciences may be even more prone to coming up with short-lived generalizations.

Does an imprecise organization theory have any scientific value? Of course. Few doubt that organizations are among the most pervasive and influential aspects of the contemporary
world (Scott and Davis, 2007). Theory about organizations need not be general, predictive, or precise to be useful. But the physical sciences are a misleading model for what organization theory should be. McCloskey (1998: 20) points out that in languages other than English, the term “science” does not carry the same baggage, as it simply means “‘disciplined inquiry,’ as distinct from, say, casual journalism or unaided common sense. It does not mean ‘quantitative,’ in the way that Lord Kelvin used it in 1883: ‘When you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.’” Disciplined inquiry can proceed even in the face of the manifest limitations I have described here.

What are the implications for research practice? Three follow directly from this analysis, for researchers, reviewers, and journals.

First, organizational researchers need to take measurement more seriously by making explicit the link between constructs and indicators, for both substantive and control variables. We have described the problems with sloppy operationalizations, such as using various weakly-correlated variables interchangeably as measures of size. Here, I believe the “social mechanisms” movement in sociology provides a good lead. Social mechanisms are the cogs and wheels of theoretical explanations, providing “an intermediary level of analysis in-between pure description and story-telling, on the one hand, and universal social laws, on the other” (Hedstrom and Swedberg, 1998). The mechanisms movement aims to dislodge the laundry-list regressions filled with endless “control” and “independent” variables that are now so readily available. Researchers should go beyond simply documenting that variables are statistically related to explain the underlying process that make it so: not just finding that a watch runs when you wind it, but prying off the back to display the cogs and wheels that make it happen (Davis and Marquis, 2005). This approach also holds at the level of measurement. In what sense are CEO tenure, or the proportion of directors that are insiders, indicators of “CEO power”? Williams et
al. (2009: 551ff.) distinguish between reflective and formative indicators, a useful distinction that is largely absent from organization-level research. If researchers were called on to explain the mechanisms presumed to link their indicators and constructs, for both substantive and control variables, much of the mischief caused by overly-abundant data would be mitigated.

Second, both researchers and reviewers should re-familiarize themselves with the validity standards that apply to quasi-experiments. The lack of experimental control need not be a fatal flaw, but the failure to recognize that a study is a quasi-experiment—and that there are better and worse ways to analyze them—often is. The list of common errors is very long, particularly in studies of performance: firms typically assign themselves to “treatments” endogenously (e.g., changing strategies or structures according to the latest fad; firing the CEO due to weak performance); treatments are not equivalent (e.g., firms putatively adopting TQM often do very different things); there is often no control group (e.g., all firms in an industry experience the same shock at the same time); and scholars often sample highly unrepresentative firms (e.g., those that went from “good to great”) in an effort to back into their “key success factors.” A refresher course in Cook and Campbell (1979) for both researchers and their reviewers would be most helpful here.

Finally, journals could help by valuing papers for their substantive importance rather than for their “contribution to theory” (e.g., by adding a new mediator or construct—cf. Colquitt and Zapata-Phelan, 2007). I agree with Hambrick (2007) that the theory benchmark prevents the publication of papers that would enrich our understanding of organizational life but do not hang on a particular paradigm, while encouraging authors to link whatever their regressions turn up to the theories currently in vogue. Creative researchers have little difficulty fitting their findings into one of the extant paradigms, along with a plausible account of what exactly the variables are measuring. Consider an analogy. The advent of the telescope in the early 17th century came at a
time when the geocentric model was still widely held in astronomy. Imagine the effect on the field if astronomers were required to show how new observations enabled by the telescope extended existing (incorrect) theory.

Here we might take note of the practices in other fields less hung up on theory, where careful attention to mechanisms and ruling out alternative interpretations takes precedence over contribution to theory. The publication norms in financial economics, for instance, allow for exceptionally brisk presentation of theory. I was exposed to these norms through writing a paper about conflicts of interest in proxy voting by mutual funds (Davis and Kim, 2007). Large mutual fund companies often run the pension funds of some of the corporations that they invest in. Efforts at corporate governance reform often take the form of shareholder-sponsored proposals at the annual meeting that are almost inevitably opposed by management. There was anecdotal evidence that fund companies frequently supported management rather than their fellow shareholders in firms whose pension funds they ran, presumably because the mutual fund companies feared being fired if they opposed management. Proxy votes had long been secret, so it was impossible to find systematic evidence on this point. Due to a new SEC requirement, however, mutual funds were required annually to disclose their proxy votes for every company in which they owned shares, starting August 31, 2004. Coupled with separate data from the Department of Labor on mutual fund companies’ pension management contracts, this allowed an answer to the question “Do mutual funds vote differently in clients than they do in non-clients?” Once the basic premise was established, no particular “theory” was required—just a careful description of the data and analyses required to answer the question, and subsidiary analyses to rule out alternative interpretations. The analyses did not particularly contribute to theory, but they did answer the question, and that was enough to count as a contribution to finance.

Organization theory has amassed an impressive armory of theoretical constructs and
mechanisms at many levels of analysis over the past half-century, and its has documented dozens of regularities. But the nature of the subject matter makes it unlikely to either yield general or precise theories. A more realistic aspiration is for carefully-done research that yields insights into particular processes at particular times. This is not intended to rule out new theory, now or in the future. It is possible to imagine new works of synthesis after a sustained period of empirical research—this was the signal contribution of Thompson’s 1967 work *Organizations in Action*. But at this stage in the field’s development, there is perhaps more value in careful empirical work on substantively important problems, even if its contribution to existing theory is minimal.
References


York: Free Press.
Figure 1: Correlations between market capitalization and employment, sales, and assets among US public corporations, 1950-2000