05-057

# The Cycles of Theory Building in Management Research

**Paul R. Carlile**
**School of Management**
**Boston University**
**Boston, MA 02215**
**carlile@bu.edu**

**Clayton M. Christensen**
**Harvard Business School**
**Boston, MA 02163**
**cchristensen@hbs.edu**

Copyright ©

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

# The Cycles of Theory Building in Management Research

Paul R. Carlile
School Of Management
Boston University
Boston, MA 02215
carlile@bu.edu

Clayton M. Christensen
Harvard Business School
Boston, MA 02163
cchristensen@hbs.edu

October 27, 2004

Version 5.0

# The *Cycle* of Theory Building in Management Research

> *Theory thus become instruments, not answers to enigmas, in which we can rest. We don't lie back upon them, we move forward, and, on occasion, make nature over again by their aid. (William James, 1907: 46)*

Some scholars of organization and strategy expend significant energy disparaging and defending various research methods. Debates about deductive versus inductive theory-building and the objectivity of information from field observation versus large-sample numerical data are dichotomies that surface frequently in our lives and those of our students. Despite this focus, some of the most respected members of our research profession (i.e., Simon (1976), Solow (1985), Staw and Sutton (1995), and Hayes (2002)) have continued to express concerns that the collective efforts of business academics have produced a paucity of theory that is intellectually rigorous, practically useful, and able to stand the tests of time and changing circumstances.

The purpose of this paper is to outline a process of theory building that links questions about data, methods and theory. We hope that this model can provide a common language about the research process that helps scholars of management spend less time defending the style of research they have chosen, and build more effectively on each other's work. Our unit of analysis is at two levels: the individual research project and the iterative cycles of theory building in which a community of scholars participates. The model synthesizes the work of others who have studied how communities of scholars cumulatively build valid and reliable theory, such as Kuhn (1962), Campbell & Stanley (1963), Glaser & Strauss (1967) and Yin (1984). It has normative and pedagogical implications for how we conduct research, evaluate the work of others, and for how we train our doctoral students.

While many feel comfortable in their own understanding of these perspectives, it has been our observation that those who have written about the research process and those who think they understand it do not yet share even a common language. The same words are applied to very different phenomena and processes, and the same phenomena can be called by many different words. Papers published in reputable journals often violate rudimentary rules for generating cumulatively improving, reliable and valid theory. While recognizing that research progress is hard to achieve at a collective level, we assert here that if scholars and practitioners of management shared a sound understanding of the process by which theory is built, we could be much more productive in doing research that doesn't just get published, but meets the standards of rigorous scholarship and helps managers know what actions will lead to the results they seek, given the circumstances in which they find themselves.

We first describe a three stage process by which researchers build theory that is at first descriptive, and ultimately normative. Second, we discuss the role that discoveries of anomalies play in the building of better theory, and describe how scholars can build theory whose validity can be verified. Finally, we suggest how scholars can define research questions, execute projects, and design student coursework that lead to the building of good theory.

# The Theory Building Process

The building of theory occurs in two major stages – the descriptive stage and the normative stage. Within each of these stages, theory builders proceed through three steps. The the theory-building process iterates through these stages again and again.[1] In the past, management researchers have quite carelessly applied the term *theory* to research activities that pertain to only one of these steps. Terms such "utility theory" in economics, and "contingency theory" in organization design, for example, actually refer only to an individual stage in the theory-building process in their respective fields. We propose that it is more useful to think of the term "theory" as a body of understanding that researchers build cumulatively as they work through each of the three steps in the descriptive and normative stages. In many ways, the term "theory" might better be framed as a verb, as much as it is a noun – because the body of understanding is continuously changing as scholars who follow this process work to improve it.

## The Building of Descriptive Theory

The descriptive stage of theory building ia a *preliminary* stage because researchers must pass through it in order to develop normative theory. Researchers who are building descriptive theory proceed through three steps: observation, categorization, and association. Terminology consistency?

### Step 1: Observation

In the first step researchers observe phenomena and carefully describe and measure what they see. Careful observation, documentation and measurement of the phenomena in words and numbers is important at this stage because if subsequent researchers cannot agree upon the descriptions of phenomena, then improving theory will prove difficult. Early management research such as *The Functions of the Executive* (Barnard, 1939) and Harvard Business School cases written in the 1940s and 50s was primarily descriptive work of this genre – and was very valuable. This stage of research is depicted in Figure 1 as the base of a pyramid because it is a necessary foundation for the work that follows. The phenomena being explored in this stage includes not just things such as people, organizations and technologies, but processes as well.

Without insightful description to subsequently build upon, researchers can find themselves optimizing misleading concepts. As an example: For years, many scholars of inventory policy and supply chain systems used the tools of operations research to derive ever-more-sophisticated optimizing algorithms for inventory replenishment. Most were based on an assumption that managers know what their levels of inventory are. Ananth Raman's pathbreaking research of the phenomena, however, obviated much of this research when he showed that most firms' computerized inventory records were broadly inaccurate – even when they used state-of-the-art automated tracking systems (Raman 199X). He and his colleagues have carefully described how inventory replenishment systems work, and what variables affect the accuracy of those processes. Having laid this foundation, supply chain scholars have now begun to build a body of theories and policies that reflect the real and different situations that managers and companies face.

[1] This model is a synthesis of models that have been developed by scholars of this process in a range of fields and scholars: Kuhn (1962) and Popper (1959) in the natural sciences; Kaplan (1964), Stinchcombe (1968), Roethlisberger (1977) Simon (1976), Kaplan (1986), Weick (1989),Eisenhardt (1989) and Van de Ven (2000) in the social sciences.

Researchers in this step often develop abstractions from the messy detail of phenomena that we term *constructs*. Constructs help us understand and visualize what the phenomena are, and how they operate. Joseph Bower's *Managing the Resource Allocation Process* (1970) is an outstanding example of this. His constructs of *impetus* and *context*, explaining how momentum builds behind certain investment proposals and fails to coalesce behind others, have helped a generation of policy and strategy researchers understand how strategic investment decisions get made. Economists' concepts of "utility" and "transactions cost" are constructs – abstractions developed to help us understand a class of phenomena they have observed. We would not label the constructs of utility and transactions cost as theories, however. They are *part* of theories – building blocks upon which bodies of understanding about consumer behavior and organizational interaction have been built.

**Step 2: Classification**

With the phenomena observed and described, researchers in the second stage then classify the phenomena into categories. In the descriptive stage of theory building, the classification schemes that scholars propose typically are defined by the attributes of the phenomena. Diversified vs. focused firms, and vertically integrated vs. specialist firms are categorization examples from the study of strategy. Publicly traded vs. privately held companies is a categorization scheme often used in research on financial performance. Such categorization schemes attempt to simplify and organize the world in ways that highlight possibly consequential relationships between the phenomena and the outcomes of interest.

Management researchers often refer to these descriptive categorization schemes as *frameworks* or *typologies*. Burgelman (1986), for example, built upon Bower's (1970) construct of context by identifying two different types of context – organizational and strategic.

**Step 3: Defining Relationships**

In the third step, researchers explore the association between the category-defining attributes and the outcomes observed. In the stage of descriptive theory building, researchers recognize and make explicit what differences in attributes, and differences in the magnitude of those attributes, *correlate* most strongly with the patterns in the outcomes of interest. Techniques such as regression analysis typically are useful in defining these correlations. Often we refer to the output of studies at this step as *models*.

Descriptive theory that quantifies the degree of correlation between the category-defining attributes of the phenomena and the outcomes of interest are generally only able to make probabilistic statements of association representing average tendencies. For example, Hutton, Miller and Skinner (2000) have examined how stock prices have responded to earnings announcements that were phrased or couched in various terms. They coded types of words and phrases in the statements as explanatory variables in a regression equation, with the ensuing change in equity price as the dependent variable. This analysis enabled the researchers then to assert that, on average across the entire sample of companies and announcements, delivering earnings announcements in a particular way would lead to the most favorable (or least unfavorable) reaction in stock price. Research such as this is important descriptive theory. However, at this point it can only assert *on average* what attributes are associated with the best

results. A specific manager of a specific company cannot know whether following that average formula will lead to the hoped-for outcome in her specific situation. The ability to know what actions will lead to desired results for a specific company in a specific situation awaits the development of normative theory in this field, as we will show below.

## The Improvement of Descriptive Theory

When researchers move from the bottom to the top of the pyramid in these three steps – observation, categorization and association – they have followed the *inductive* portion of the theory building process. Theory begins to improve when researchers cycle from the top back to the bottom of this pyramid in the *deductive* portion of the cycle – seeking to "test" the hypothesis that had been inductively formulated. This most often is done by exploring whether the same correlations exist between attributes and outcomes in a different set of data than the data from which the hypothesized relationships were induced. When scholars test a theory on a new data set (whether the data are numbers in a computer, or are field observations taken in a new context), they might find that the attributes of the phenomena in the new data do indeed correlate with the outcomes as predicted. When this happens, this "test" confirms that the theory is of use under the conditions or circumstances observed.[2] However, the researcher returns the model to its place atop the pyramid *tested but unimproved*.

It is only when an anomaly is identified – an outcome for which the theory can't account – that an opportunity to improve theory occurs. As Figure 1 suggests, discovery of an anomaly gives researchers the opportunity to revisit the categorization scheme – to cut the data in a different way – so that the anomaly *and* the prior associations of attributes and outcomes can all be explained. In the study of how technological innovation affects the fortunes of leading firms, for example, the initial attribute-based categorization scheme was radical vs. incremental innovation. The statements of association that were built upon it concluded that the leading established firms on average do well when faced with incremental innovation, but they stumble in the face of radical change. But there were anomalies to this generalization – established firms that successfully implemented radical technology change. To account for these anomalies, Tushman & Anderson (1986) offered a different categorization scheme, competence-enhancing vs. competence-destroying technological changes. This scheme resolved many of the anomalies to the prior scheme, but subsequent researchers uncovered new ones for which the Tushman-Anderson scheme could not account. Henderson & Clark's (1990) categories of modular vs. architectural innovations; Christensen's (1997) categories of sustaining vs. disruptive technologies; and Gilbert's (2001) threat-vs.-opportunity framing each uncovered and resolved anomalies for which the work of prior scholars could not account. This body of understanding has improved and become remarkably useful to practitioners and subsequent scholars (Adner, 2003; Daneels, 2005) *because* these scholars followed the process in a disciplined way: – uncovered anomalies, sliced the phenomena in different ways, and articulated new associations between the attributes that defined the categories and the outcome of interest.

[2] Popper asserts that a researcher in this phase, when the theory accurately predicted what he observed, can only state that his test or experiment of the theory "corroborated" or "failed to dis-confirm" the theory.

Figure 1

**The Process of Building Theory**

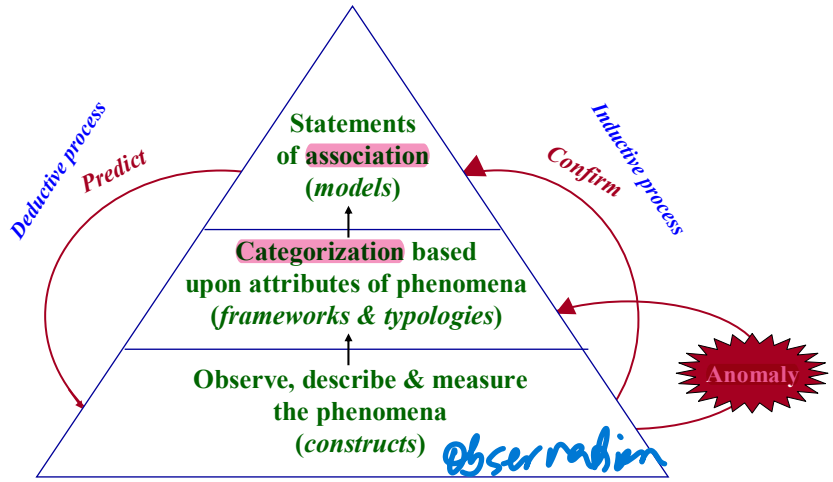



Figure 1 suggests that there are two sides to every lap around the theory-building pyramid: an inductive side and a deductive side. In contrast to either/or debates about the virtues of deductive and inductive approaches to theory, this suggests that any complete cycle of theory building includes both.[3]

Descriptive theory-building efforts typically categorize by the attributes of the phenomena because they are easiest to observe and measure. Likewise, correlations between attributes and outcomes are easiest to hypothesize and quantify through techniques such as regression analysis. Kuhn (1962) observed that confusion and contradiction typically are the norm during descriptive theory-building. This phase is often characterized by a plethora of categorization schemes, as in the sequence of studies of technology change cited above, because the phenomena generally have many different attributes. Often, no model is irrefutably superior: Each seems able to explain anomalies to other models, but suffers from anomalies to its own.

## The Transition from Descriptive to Normative Theory

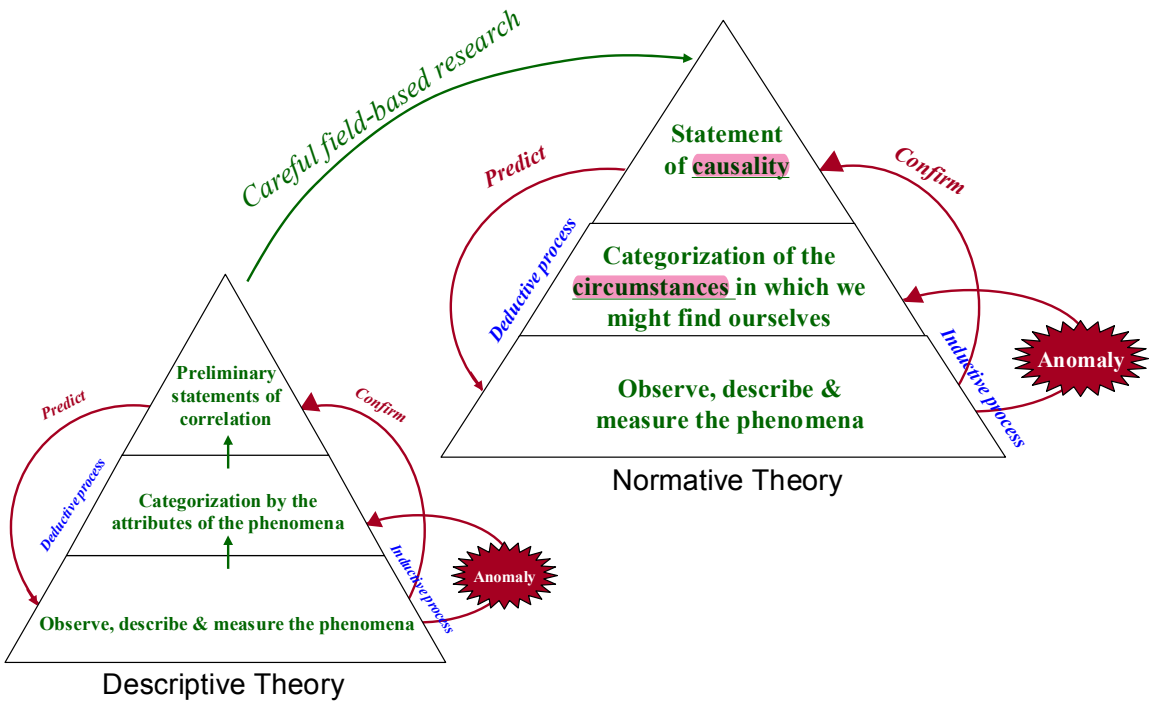
The confusion and contradiction that often accompany descriptive theory become resolved when careful researchers – often through detailed empirical and ethnographic observation – move beyond statements of correlation to define what *causes* the outcome of interest. As depicted in Figure 2, they leap across to the top of the pyramid of causal theory. With their understanding of causality, researchers then work to improve theory by following the same three steps that were

[3] Kant, Popper, Feyerabend and others have noted that all observations are shaped, consciously or unconsciously, by cognitive structures, previous experience or some theory-in-use. While it is true that individual researchers might start their work at the top of the pyramid, we believe that the hypotheses that deductive theorists test generally had been derived consciously or unconsciously, by themselves or others, from an inductive source. There are few blue-sky hypotheses that were formulated in the complete absence of observation.

used in the descriptive stage. Hypothesizing that their statement of causality is correct, they cycle deductively to the bottom of the pyramid to test the causal statement: If we observe these actions being taken, these should be the outcomes that we observe. When they encounter an anomaly, they then delve into the categorization stage. Rather than using schemes based on attributes of the phenomena, however, they develop categories of the different situations or circumstances in which managers might find themselves. They do this by asking, when they encounter an anomaly, "What was it about the situation in which those managers found themselves, that caused the causal mechanism to yield a different result? By cycling up and down the pyramid of normative theory, researchers will ultimately define the set of the situations or circumstances in which managers might find themselves when pursuing the outcomes of interest. This allows researchers to make contingent statements of causality – to show how and why the casual mechanism results in a different outcome, in the different situations. A theory completes the transition from descriptive to normative when it can give a manager unambiguous guidance about what actions will and will not lead to the desired result, given the circumstance in which she finds herself.

Figure 2:

The Transition from Descriptive Theory to Normative Theory




The history of research into manned flight is a good way to visualize how this transition from descriptive to normative theory occurs, and how it is valuable. During the middle ages, would-be aviators did their equivalent of best-practices research and statistical analysis. They observed the many animals that could fly well, and compared them with those that could not. The vast majority of the successful fliers had wings with feathers on them; and most of those that couldn't fly had neither. This was quintessential descriptive theory. Pesky outliers like ostriches

had feathered wings but couldn't fly; bats had wings without feathers and were very good at it; and flying squirrels had neither and got by. But the $R^2$ was so high that aviators of the time copied the seemingly salient characteristics of the successful fliers in the belief that the visible attributes of the phenomena *caused* the outcome. They fabricated wings, glued feathers on them, jumped off cathedral spires, and flapped hard. It never worked. For centuries they assumed that the prior aviators had failed because they had bad wing designs; hadn't bulked up their muscles enough; or hadn't flapped hard enough. There were substantial disagreements about which of the birds' attributes truly enabled flight. For example, Roger Bacon in about 1285 wrote an influential paper asserting that the differentiating attribute was birds' hollow bones (Clegg, 2003). Because man had solid bones, Bacon reasoned, we could never fly. He then proposed several machine designs that could flap their wings with sufficient power to overcome the disadvantage of solid bones. But it still never worked. Armed with the correlative statements of descriptive theory, aviators kept killing themselves.

Then through his careful study of fluid dynamics Daniel Bernoulli identified a shape that we call an airfoil – a shape that, when it cuts through air, creates a mechanism that we call lift. Understanding this causal mechanism, which we call Bernoulli's Principle, made flight possible. But it was not yet predictable. In the language of this paper, the theory predicted that aviators would fly successfully when they built machines with airfoils to harness lift. But while they sometimes flew successfully, occasionally they did not. Crashes were anomalies that Bernoulli's theory could not explain. Discovery of these anomalies, however, allowed the researchers to revisit the categorization scheme. But this time, instead of slicing up the world by the attributes of the good and bad fliers, researchers categorized their world by circumstance – asking the question, "What was it about the circumstance that the aviator found himself in that caused the crash?" This then enabled them to improve equipment and techniques that were based upon circumstance-contingent statements of causality: "This is how you should normally fly the plane. But when you get in this situation, you need to fly it differently in order to get the desired outcome. And when you get in that situation, don't even try to fly. It is impossible."

When their careful studies of anomalies allowed researchers to identify the set of circumstances in which aviators might find themselves, and then modified the equipment or developed piloting techniques that were appropriate to each circumstance, manned flight became not only possible, but predictable. Hence, it was the discovery of the fundamental *causal* mechanism that made flight possible. And it was the categorization of the salient circumstances that made flight predictable. This is how this body of understanding about human flight transitioned from descriptive to normative theory.

Dsciplined scholars can achieve the same transition in management research. The discovery of the fundamental causal mechanisms makes it *possible* for managers purposefully to pursue desired outcomes successfully and predictably. When researchers categorize managers' world according to the circumstances in which they might find themselves, they can make circumstance-contingent statements of cause and effect, of action and result.

**Circumstance-based categories and normative theory**

Some cynical colleagues despair of any quest to develop management theories that make success possible and predictable – asserting that managers' world is so complex that there are an

infinite number of situations in which they might find themselves. Indeed, this is very nearly true in the descriptive theory phase. But normative theory generally is not so confusing. Researchers in the normative theory phase resolve confusion by abstracting up from the detail to define a few categories – typically two to four – that comprise *salient* circumstances. Which boundaries between circumstances are salient, and which are not? Returning to our account of aviation research, the boundaries that defined the salient categories of circumstance are determined by the necessity to pilot the plane differently. If a different circumstance does not require different methods of piloting, then it is not a meaningful category. The same principle defines the salience of category boundaries in management theory. If managers find themselves in a circumstance where they must change actions or organization in order to achieve the outcome of interest, then they have crossed a salient boundary.

Several prominent scholars have examined the improvement in predictability that accompanies the transition from the attribute-based categorization of descriptive theory, to the circumstance-based categorization of normative theory. Consider, for example, the term "Contingency Theory" – a concept born of Lawrence & Lorsch's (1967) seminal work. They showed that the best way to organize a company depended upon the circumstances in which the company was operating. In our language, contingency is not a theory *per se*. Rather, contingency is a crucial element of *every* normative theory – it is the categorization scheme. Rarely do we find one-size-fits-all answers to every company's problem. The effective course of action will generally "depend" on the circumstance.

Glaser and Strauss's (1967) treatise on "grounded theory" actually is a book about categorization. Their term *substantive theory* corresponds to the attribute-defined categories in *descriptive theory*. And their concept of *formal* theory matches our definition of normative theory that employs categories of circumstance..

Thomas Kuhn (1962) discussed in detail the transition of understanding from descriptive to normative theory in his study of the emergence of scientific paradigms. He described a preliminary period of confusion and debate in theory building, which is an era of descriptive theory. His description of the emergence of a paradigm corresponds to the transition to normative theory described above. We agree with Kuhn that even when a normative theory achieves the status of a broadly believed paradigm, it continues to be improved through the process of discovering anomalies, as we describe above. Indeed, the emergence of new phenomena – which probably happens more frequently in competitive, organizational and social systems than in the natural sciences – ensures that there will always be additional productive laps up and down the theory pyramid that anomaly-seeking researchers can run.

The observation that management research is often faddish has been raised enough that it no longer seems shocking (Micklethwait and Wooldridge, 1996; Abrahamson, 1998). Fads come and go when a researcher studies a few successful companies, finds that they share certain characteristics, concludes that he has seen enough, and then skips the categorization step entirely by writing a book asserting that if all managers would imbue their companies with those same characteristics, they would be similarly successful. When managers then apply the formula and find that it doesn't work, it casts a pall on the idea. Some faddish theories aren't uniformly bad. It's just that their authors were so eager for their theory to apply to everyone that they never took the care to distinguish correlation from causality, or to figure out the circumstances in which their

statement of causality would lead to success, and when it would not. Efforts to study and copy "the best practices of successful companies" almost uniformly suffer from this problem.

Unfortunately, it is not just authors-for-profit of management books that contribute to the problem of publishing theory whose application is uncertain. Many academics contribute to the problem by taking the other extreme – articulating tight "boundary conditions" outside of which they claim nothing. Delimiting the applicability of a theory to the specific time, place, industry and/or companies from which the conclusions were drawn in the first place is a mutation of one of the cardinal sins of research – sampling on the dependent variable. In order to be useful to managers and to future scholars, researchers need to help managers understand the circumstance that they are in. Almost always, this requires that they also be told about the circumstances that they are *not* in.

## The Value of Anomalies

As indicated before, when researchers in both the descriptive and normative stages use statements of association or causality to predict what they will see, they often observe something that the theory did *not* lead them to expect; thus identifying an *anomaly*—something the theory could not explain. This discovery forces theory builders to cycle back into the categorization stage with a puzzle such as "there's something else going on here" or "these two things that we thought were different, really aren't." The results of this effort typically can include: 1) more accurately describing and measuring what the phenomena are and are not; 2) changing the definitions by which the phenomena or the circumstances are categorized – adding or eliminating categories or defining them in different ways; and/or 3) articulating a new theoretical statement of what is associated with, or causes what, and why, and under what circumstances. The objective of this process is to revise theory so that it still accounts for both the anomalies identified and the phenomena as previously explained.

Anomalies are valuable in theory building because *the discovery of an anomaly is the enabling step to identifying and improving the categorization scheme in a body of theory* – which is the key to being able to apply the theory with predictable results. Researchers whose goal is to "prove" a theory's validity are likely to view discovery of an anomaly as failure. Too often they find reasons to exclude outlying data points in order to get more significant measures of statistical fit. There typically is more information in the points of outlying data than in the ones that fit the model well, however, because understanding the outliers or anomalies is generally the key to discovering a new categorization scheme. This means that journal editors and peer reviewers whose objective is to improve theory should embrace papers that seek to surface and resolve anomalies.

Indeed, productive theory-building research is almost invariably prompted or instigated by an anomaly or a paradox (Poole & Van de Ven, 1989). The research that led to Michael Porter's (1991) *Competitive Advantage of Nations* is an example. Before Porter's work, the theory of international trade was built around the notion of comparative advantage. Nations with inexpensive electric power, for example, would have a competitive advantage in those products in which the cost of energy was high; those with low labor costs would enjoy an advantage in making and selling products with high labor content; and so on. Porter saw anomalies for which this theory could not account. Japan, with little iron ore and coal, became a successful steel

producer. Italy became the world's dominant producer of ceramic tile even though it had high electricity costs and had to import much of the clay used in making the tile. Porter's work categorized the world into two circumstances – situations in which a factor-based advantage exists, and those in which it does not. In the first situation the reigning theory of comparative advantage still has predictive power. But in the latter circumstance, Porter's theory of competitive industrial clusters explained the phenomena that had been anomalous to the prior theory. Porter's theory is normative because it gives planners clear guidance about what they should do, given the circumstance in which they find themselves. The government of Singapore, for example, attributes much of that country's prosperity to the guidance that Porter's theory has provided.

Yin (1984) distinguishes between *literal* replications of a theory, versus *theoretical* replications. A literal replication occurs when the predicted outcome is observed. A theoretical replication occurs when an unusual outcome occurs, but for reasons that can be explained by the model. Some reviewers cite "exceptions" to a theory's predictions as evidence that it is invalid. We prefer to avoid using the word "exception" because of its imprecision. For example, the observation that airplanes fly is an exception to the general assertion that the earth's mass draws things down toward its core. Does this exception disprove the theory of gravity? Of course not. While falling apples and flaming meteors are literal replications of the theory, manned flight is a theoretical replication. It is a different outcome than we normally would expect, but Bernoulli's Principle explains why. An anomaly is an outcome that is neither a literal or theoretical replication of a theory.

# How to Design Anomaly-Seeking Research

Although some productive anomalies might be obvious from the outset, often the task of theory-building scholars is to design their research to *maximize the probability that they will be able to identify anomalies*. Here we describe how to define research questions that focus on anomalies, and outline three ways to design anomaly-seeking research. We conclude this section by describing how literature reviews might be structured to help readers understand how knowledge has accumulated in the past, and position the present paper in the stream of scholarship.

## Anomaly-Seeking Research Questions

Anomaly-seeking research enables new generations of researchers to pick up even well-accepted theories, and to run the theory-building cycle again – adding value to research that already has earned broad praise and acceptance. Consider Professor Porter's (1991) research mentioned above. In Akron, Ohio there was a powerful cluster of tire manufacturers whose etiologies and interactions could be explained well by Porter's theory. That group subsequently vaporized – in part because of the actions of a company, Michelin, that operated outside of this cluster (Sull, 2000). This anomaly suggests that there must situations in time or space in which competing within a cluster is competitively important; in other situations it must be less important. When an improved categorization scheme emerges from Sull's and others' work, the community of scholars and policy makers will have an even clearer sense for when the competitive crucible of clusters is critical for developing capabilities, when it is not, and why.

In this spirit, we outline below some examples of "productive" questions that could be pursed by future researchers that potentially challenge many current categories used in management research:

- When might process re-engineering or lean manufacturing be *bad* ideas?
- When could sourcing from a partner or supplier something that is *not* your core competence lead to *disaster*?
- Are there circumstances in which pencil-on-paper methods of vendor management yield better results than using supply-chain management software?
- When and why is a one-stop-shopping or "portal" strategy effective and when would we expect firms using focused specialist strategies to gain the upper hand?
- When are time-based competition and mass customization likely to be critical and when might they be competitively meaningless?
- Are SIC codes the right categories for defining "relatedness" in diversification research?
- When should acquiring companies integrate a firm they have just purchased into the parent organization, and when should they keep it separate?

Much published management research is of the half-cycle, terminal variety – hypotheses are defined and "tested." Anomaly-seeking research *always* is focused on the categorization step in the pyramid. Many category boundaries (such as SIC codes) seem to be defined by the availability of data, rather than their salience to the underlying phenomena or their relation to the outcome – and questioning their sufficiency is almost always a productive path for building better theory. "When doesn't this work?" and "Under what conditions might this gospel be bad news?" are simple questions that can yield breakthrough insights – and yet too few researchers have the instinct to ask them.

## The Lenses of Other Disciplines

One of Kuhn's (1962) most memorable observations was that the anomalies that led to the toppling of a reigning theory or paradigm almost invariably were observed by researchers whose backgrounds were in different disciplines than those comprising the traditional training of the leaders in the field. The beliefs that adherents to the prior theory held about what was and was not possible seemed to shape so powerfully what they could and could not see that they often went to their graves denying the existence or relevance of the very anomalous phenomena that led to the creation of improved theory. Researchers from different disciplines generally use different methods and have different interests toward their object of study. Such differences often allow them to see things that might not be recognized or might appear inconsequential to an insider.

It is not surprising, therefore, that many of the most important pieces of breakthrough research in the study of management, organization and markets have come from scholars who stood astride two or more academic disciplines. Porter's (1980, 1985, 1991) work in strategy, for

example, resulted from his having combined insights from business policy and industrial organization economics. The insights that Robert Hayes and his colleagues (1980, 1984, 1985, 1988) derived about operations management combined insights from process research, strategy, cost accounting and organizational behavior. Baldwin & Clark's (2000) insights about modularity were born at the intersection of options theory in finance with studies of product development.

Clark Gilbert ((2001) looked at Christensen's (1997) theory of disruptive innovation through the lenses of prospect theory and risk framing (Kahnemann & Tversky 1979, 1984), and saw explanations of what had seemed to be anomalous behavior, for which Christensen's model could not account.

## Studying the Phenomena within the Phenomena

The second method to increase the probability that researchers will identify anomalies is to execute *nested* research designs that examine different levels of phenomena. Rather than study just industries *or* companies *or* divisions *or* groups *or* individuals, a nested research design entails studying how individuals act and interact within groups; and how the interaction amongst groups and the companies within which they are embedded affect the actions of individuals. Many anomalies will only surface while studying second-order interactions across levels within a nested design.

The research reported in Johnson & Kaplan's *Relevance Lost* (1987) which led to the concept of activity-based costing, is a remarkable example of the insights gained through nested research designs. Most prior researchers in managerial accounting and control had conducted their research at a single level—the numbers printed in companies' financial statements. Johnson and Kaplan saw that nested beneath each of those printed numbers was a labyrinth of political, negotiated, judgmental processes that could systematically yield inaccurate numbers.

Spear and Bowen (1999) developed their path-breaking insights of the Toyota Production System through a nested research design. Researchers in the theory's descriptive stage had studied Toyota's production system at single levels. They documented visible artifacts such as minimal inventories, *kanban* scheduling cards and rapid tool changeovers. After comparing the performance of factories that did and did not possess these attributes, early researchers asserted that if other companies would use these same tools, they could achieve similar results (see, for example, Womack *et.al.*, 1990). The anomaly that gripped Spear and Bowen was that when other firms used these artifacts, they still weren't able to achieve Toyota's levels of efficiency and improvement. By crawling inside to study how individuals interacted with individuals, in the context of groups interacting with other groups, within and across plants within the company and across companies, Spear and Bowen were able to go beyond the correlative statements of descriptive theory, to articulate the fundamental causal mechanism behind the Toyota system's self-improving processes – which they codified as four "rules-in-use" that are not written anywhere but are assiduously followed when designing processes of all sorts at Toyota.

Spear is now engaged in search of anomalies on the deductive side of the cycle of building normative theory. Because no company besides Toyota has employed this causal mechanism, Spear cannot retrospectively study other companies. Like Johnson & Kaplan did when they used

How people work

Rule #1: All work shall be highly specified as to content, sequence, timing and outcome.

How people connect

Rule #2: Every customer-supplier connection must be direct, and there be an unambiguous yes-or-no way to send requests and receive responses.

How the production line is contructed Rule #3: The pathway for every product and service must be simple and direct.

How to improve Rule #4: Any improvement must be made in accordance with the scientific method, under the guidance of a teacher, at the lowest possible in the organization.

"action research" to study the implementation problems of activity-based costing, Spear is helping companies in very different circumstances to use his statements of causality, to see whether the mechanism of these four rules yields the same results. To date, companies in industries as diverse as aluminum smelting, hospitals, and jet engine design have achieved the results that Spear's theory predicts – he has not yet succeeded in finding an anomaly. The categorization step of this body of normative theory still has no salient boundaries within it.

## Observing and Comparing a Broad Range of Phenomena

The third mechanism for maximizing the probability of surfacing an anomaly is to examine, in the deductive half of the cycle, a broader range of phenomena than prior scholars have done. As an example, Chesbrough's (1999) examination of Japanese disk drive makers (which Christensen had excluded from his study) enabled Chesbrough to surface anomalies for which Christensen's theory of disruptive technology could not account—leading to an even better theory that then explains a broader range of phenomena. The broader the range of outcomes, attributes and circumstances that are studied at the base of the pyramid, the higher the probability that researchers will identify the salient boundaries among the categories.

## Anomaly-Seeking Research and the Cumulative Structure of Knowledge

When interviewing new faculty candidates who have been trained in methods of modeling, data collection and analysis as doctoral students, we observe that many seem almost disinterested in the value of *questions* that their specialized techniques are purporting to answer. When asked to position their work upon a stream of scholarship, they recite long lists of articles in "the literature," but then struggle when asked to diagram within that body of work which scholar's work resovles anomalies to prior scholars' theories; whose results contradicted whose, and why. Most of these lists of prior publications are simply lists, sometimes lifted from prior authors' lists of prior articles. They are listed because of their relatedness to the topic. Few researchers have been taught to organize citations in a way that describes the laps that prior researchers have taken, to give readers a sense for how theory has or has not been built to date. Rather, after doffing the obligatory cap to prior research, they get busy testing their hypotheses in the belief that if nobody has tested these particular ones before, using novel analytical methods on a new data set, it breaks new ground.

Our suggestion is that in the selection of research questions and the design of research methods, authors physically map the literature on a large sheet of paper in the format of Figure 2 above, and then answer questions like these:

- Is this body of theory in the descriptive or normative stage?
- What anomalies have surfaced in prior authors' work, and which pieces of research built on those by resolving the anomaly? In this process, how have the categorization schemes in this field improved?
- At what step am I positioning my work? Am I at the base of the pyramid defining constructs to help others abstract from the detail of the phenomena what really is going on? Am I strengthening the foundation by offering better ways to examine and measure

the phenomena more accurately? Am I resolving an anomaly by suggesting that prior scholars haven't categorized things correctly? Am I running half a lap or a complete cycle, and why?

Similarly, in the "suggestions for future research" section of the paper, we suggest that scholars be much more specific about where future anomalies might be buried. "Who should pick up the baton that I am setting down at the end of my lap, and in what direction should they run?"

We have attempted to construct such maps in several streams of research with which we are familiar (See, for example, Gilbert 2005). It has been shocking to see how difficult it is to map how knowledge has accumulated within a given sub-field. In many cases, it simply hasn't summed up to much, as the critics cited in our first paragraph have observed.

We suggest that the pyramids of theory building might constitute a generic map, of sorts, to bring organization to the collective enterprises within each field and sub-field. The curriculum of doctoral seminars might be organized in this manner, so that students are brought through the past into the present in ways that help them visualize the next steps required to build better theory. Literature reviews, if constructed in this way at the beginning of papers, would help readers position the work in the context of this stream, in a way that adds much more value than listing articles that are topically related.

Here's just one example of how this might be done. Alfred Chandler's (1977, 1990) landmark studies essentially proposed a theory: that the "visible hand" of managerial capitalism was a crucial enabling factor that led not just to rapid economic growth between 1880 and 1930, but led to the dominance of industry after industry by large, integrated corporations that had the scale and scope to pull everything together. In recent years, much has been written about "virtual" corporations and "vertical dis-integration;" indeed, some of today's most successful companies such as Dell are specialists in just one or two slices of the vertical value-added chain. To our knowledge, few of the studies that focus on these new virtual forms of industrial organization have even hinted that the phenomena they are focusing upon actually is an anomaly for which Chandler's theory of capitalism's visible hand cannot adequately account. If these researchers were to build their work on this anomaly, it would cause them to delve back into the categorization process. Such an effort would define the circumstances in which technological and managerial integration of the sort that Chandler observed are crucial to building companies and industries, while identifying other circumstances in which specialization and market-based coordination are superior structures. A researcher who structured his or her literature review around this puzzle, and then executed that research, would give us a better contingent understanding of what causes what and why.

## Establishing the Validity of Theory

A primary concern of every consumer of management theory is to understand where it applies, and where it does not apply. Yin (1984) helps us with these concerns by defining two types of validity for a theory – internal and external validity – which are the dimensions of a body of understanding that help us guage whether and when we can trust it. In this section we'll discuss how these concepts relate to our model of theory building, and describe how researchers can make their theories valid on both of these dimensions.

## Internal Validity

Yin asserts that a theory's *internal* validity is the extent to which: 1) its conclusions are logically drawn from its premises; and 2) the researchers have ruled out all plausible alternative explanations that might link the phenomena with the outcomes of interest. The best way we know to ensure the internal validity of a theory is to examine the phenomena through the lenses of as many disciplines and parts of the company as possible – because the plausible alternative explanations almost always are found in the workings of another part of the company, as viewed through the lenses of other academic disciplines. We offer here two illustrations.

Intel engineered a remarkable about-face in the early 1980s, as it exited the industry it built – Dynamic Random Access Memories (DRAMs) – and threw all of its resources behind its microprocessor strategy. Most accounts of this impressive achievement attribute its success to the leadership and actions of its visionary leaders, Gordon Moore and Andy Grove (see, for example, Yoffie *et.al.* 2002). Burgelman's careful ethnographic reconstruction of the resource allocation process within Intel during those years of transition, however, reveals a very different explanation of how and why Intel was able to make this transition. As he and Grove have shown, it had little to do with the decisions of the senior-most management (Burgelman, 2002).

One of the most famous examples of research that strengthens its internal validity by examining a phenomenon through the lenses of several disciplines is Graham Allison's (1971) *The Essence of Decision*. Allison examined the phenomena in a single situation—the Cuban missile crisis—using the assumptions of three different theoretical lenses (e.g., rational actor, organizational, & bureaucratic). He surfaced anomalies in the current understanding of decision making that could not have been seen had he only studied the phenomenon from a single disciplinary perspective. Through the use of multiple lenses he contributed significantly to our understanding of decision making in bureaucratic organizations.

As long as there's the possibility that another researcher could say, "Wait a minute. There's a totally different explanation for why this happened," then we cannot be assured of a theory's internal validity. If scholars will patiently examine the phenomena and outcomes of interest through the lenses of these different perspectives, they can incorporate what they learn into their explanations of causality. And one-by-one, they can rule out other explanations so that theirs is the only plausible one left standing. It can then be judged to be internally valid.

## External Validity

The *external* validity of a theory is the extent to which a relationship that was observed between phenomena and outcomes in one context can be trusted to apply in different contexts as well. Many researchers have come to believe that a theory's external validity is established by "testing" it on different data sets. This can *never* conclusively establish external validity, however – for two reasons. First, researchers cannot test a theory on every conceivable data set; and second, data only exists about the past. How can we be sure a model applies in the future, when there is no data to test it on? Consider, for illustration, Christensen's experience after publishing the theory of disruptive innovation in *The Innovator's Dilemma* (Christensen, 1997). This book presented in its first two chapters a normative theory, built upon careful empirical descriptions of the history of the disk drive industry. It asserted that there are two circumstances

– sustaining and disruptive situations – in which innovating managers might find themselves. Then it defined a causal mechanism – the functioning of the resource allocation process in response to the demands of customers and financial markets – that caused leading incumbent firms and entrants to succeed or fail at different types of innovations in those circumstances.

Christensen's early papers summarized the history of innovation in the disk drive industry, from which the theory was inductively derived. Those who read these papers instinctively wondered, "Does this apply outside the disk drive industry?" In writing *The Innovator's Dilemma*, Christensen sought to establish the generalizability or external validity of the theory by "testing" it on data from as disparate a set of industries as possible – including hydraulic excavators, steel, department stores, computers, motorcycles, diabetes care, accounting software, motor controls and electric vehicles. Despite the variety of industries in which the theory seemed to have explanatory power, executives from industries that weren't specifically studied kept asking, "Does it apply to health care? Education? Financial services?" When Christensen published additional papers that applied the model to these industries, the response was, "Does it apply to telecommunications? Relational database software? Does it apply to Germany" The killer question, from an engineer in the disk drive industry, was, "It clearly applies to the *history* of the disk drive industry. But does it apply to its *future* as well? Things are very different now." As these queries illustrate, it is simply impossible to establish the external validity of a theory by testing it on data sets – because there will always be another one upon which it hasn't yet been tested, and the future will always lie just beyond the reach of data.

When researchers have defined what causes what, and why, and show how the result of that causal mechanism differs by circumstance, then the scope of the theory, or its external validity, is established. In the limit, we could only say that a theory is externally valid when the process of seeking and resolving anomaly after anomaly results in a set of categories that are collectively exhaustive and mutually exclusive. Mutually exclusive categorization would allow managers to say, "I am in this circumstance and not that one." And collectively exhaustive categorization would assure us that all situations in which managers might find themselves with respect to the phenomena and outcomes of interest, are accounted for in the theory. No theory's categorization is likely to achieve the ultimate status of mutually exclusive and collectively exhaustive, of course. But the accumulation of insights and improvements from cycles of anomaly-seeking research can improve theory asymptotically towards that goal.

This raises an interesting paradox for large sample-size research that employs "mean" analyses to understand ways to achieve the optimum result or best performance. One would think that a theory derived from a large data set representing an entire population of companies would have greater external validity than a theory derived from case studies of a limited number of situations within that population. However, when the unit of analysis is a population of companies, the researcher can be specific only about the entire population of companies – the population comprises one category, and other sources of variance or differences that exist in that population become potentially lost as an explanation. Some managers will find that following the formula that works best on average, works best in their situation as well, of course. However, sometimes the course of action that is optimal on average will *not* yield the best outcome in a specific situation. Hence, researchers who derive a theory from statistics about a population still need to establish external validity through circumstance-based categorization.

Some large sample, quantitative studies in strategy research have begun to turn to analyses that estimate simultaneously the expected value (a mean analysis) and the variance associated with performance oriented dependent variables using a "variance decomposition" approach (Fleming and Sorensen, 2001; Sorensen and Sorensen, 2001). The simultaneous nature of this methodological approach allows a deeper understanding of the mean as well as the variance associated with a firm overtime (Sorensen, 2002) or a population of firms (Hunter, 2002). What such analysis suggests is that when there are significant heterogeneity in a given strategic environment, not only will there be variance in firm performance, but also what a firm needs to do to be successful will also differ based of the niche that they pursue. This reminds us that explanations for strategic questions are not only contingent, but more importantly are based on an understanding what sources of variance, what relations across different variables, matter most and why. From a methodological point of view, this also reminds of how our abilities (i.e., tools, methods) to represent data shape how we are able to describe what "strategic action" is possible.

The value of progressing from descriptive to normative theory can be illustrated in the case of Jim Collins' (2001) popular book, *Good to Great*. Collins and his research team found 15 companies that had gone from a period of mediocre performance to a period of strong performance. They then found a matching set of companies in similar industries that had gone from mediocre performance to another period of mediocre performance, and identified attributes that the "good-to-great" companies shared in common, and found that the "good-to-good" companies did not share these attributes. Greater success is associated with the companies that possess these attributes. They have done a powerful piece of descriptive theory-building built on a categorization scheme of companies that share these attributes, vs. companies that do not.

The research in this book has been very helpful to *many* executives and academics. As descriptive theory, however, there is still uncertainty about whether a specific company in a specific situation will succeed if it acquires the attributes of the good-to-great, because the theory has not yet gone through the process of circumstance-based categorization. For example, one of those attributes is that the good-to-great companies were led by relatively humble CEOs who generally have shunned the limelight, whereas the mediocre companies tended to be led by more ego-centric, hired-in "superstar" executives. There might indeed be situations in which an ego-centric superstar executive is crucial to success, however. Such a precise, situation-specific statement will only possible – and the theory can be judged to be externally valid – only when this body of understanding has progressed to the normative theory stage.

## What is Good Data?

The dichotomy between subjectivity and objectivity is often used as a cleavage point to judge the scientific quality of data – with many seeing objective data as more legitimate than subjective data. Case- or field-derived data versus large-sample data sets is a parallel dichotomy that often surfaces in academic discourse. Much like theory, the only way we can judge the value of data is by their usefulness in helping us understand how the world works, identifying categories, making predictions and surfacing anomalies.

Research that employs a nested design often reveals how illogical these dichotomies are. Christensen's (1997) research, for example, was built upon a history of the disk drive industry derived from analysis of tens of thousands of data points about markets, technologies and

products that were reported in *Electronic Business* and *Disk/Trend Report*. In the context of the industry's history, the study then recounted the histories of individual companies, which were assembled partially from published statistics and partially from interviews with company managers. The study also included histories of product development projects within these companies, based upon a few numbers and extensive personal interviews. Finally, the study included many accounts of individuals' experiences in developing and launching new products, comprised exclusively of information drawn from interviews – with no numbers included whatsoever. So what is a case study? Because a case is a description and assessment of a situation over a defined period of time, every level in Christensen's study was a case – industry, company, group and individual.

And what is data? Each level of this study involved lots of data of many sorts. Each of these descriptions – from the industry's history to the individuals' histories – captured but a fraction of the richness in each of the situations. Indeed, the "hardest" numbers on product product performance, company revenues and competitors' market shares, really were after-the-fact proxy manifestations all the processes, prioritizations and decisions amongst the groups and individuals that were observed in the nested, "subjective" portions of the study.

Let's drill more deeply on this question of where much quantitative data comes from. For example, the data used in many research projects comes directly or indirectly from the reported financial statements of publicly traded companies. Is this objective data? Johnson & Kaplan (1987) showed quite convincingly that the numbers representing revenues, costs and profits that appear in companies' financial statements are typically the result of processes of estimation, allocation, debate and politics that can produce grossly inaccurate reflections of true cost and profit. The subjective nature of financial statement data, and the skills and methods used by those who made those judgments, however, are hidden from the view of researchers who use the published numbers.

The healthiest and probably the most accurate mindset for researchers is that nearly all research – whether presented in the form of large data sample analysis, a mathematical optimization model, or an ethnographic description of behavior – is a description of a situation and is, therefore, a case. And *all data are subjective*. Each form of data is a higher-level abstraction from a much more complex reality, out of which the researcher attempts to pull the most salient variables or patterns for examination. Generally, the subjectivity of data is glaringly apparent in field-based, ethnographic research, whereas the subjectivity tends to be hidden behind numerical data.

Researchers of every persuasion ought always to strive to examine phenomena not just through the lenses of different academic or functional disciplines, but through the lenses of multiple forms of data as well. And none of us ought to be defensive or offensive about the extent to which the data in our or others' research are subjective. We are all in the same boat, and are obligated to do our best to be humble and honest with ourselves and our colleagues as we participate individually within and collectively across the theory building cycle.[4]

---

[4] An excellent account that has helped us understand how pervasive the exercise of subjectivity is in the generation of "facts" is E.H. Carr's (1961) treatise, *What Is History*. Carr describes that even the most complete historical accounts simply summarize what those who recorded events decided were important or interesting enough to record. In most

## Implications for Course Design

Schools of management generally employ two methods of classroom instruction: case-based classes and lecture-based classes. These are descriptive categorizations of the phenomena. Attempts to assess which method of instruction is associated with the best outcomes is fraught with anomaly. We suggest that there is a different, circumstance-based categorization scheme that may constitute a better foundation of a theory of course design: Whether the instructor is using the course to *develop* theory, or to help students practice the *use* of theory.

When designing a course on a subject about which normative theory has not yet emerged, designing the course to move up the *inductive* side of the theory pyramid can be very productive. For example, Harvard Business School professor Kent Bowen decided several years ago that because a significant portion of HBS graduates end up running small businesses, he ought to create a course that prepares students to do that. He then discovered that the academic literature was amply stocked with studies of how to structure deals and start companies, but that there wasn't much written about how to run plain old low-tech, slow-growth companies. Bowen tackled the problem with an inductive course-design strategy. He first wrote a series of cases that simply described what managers in these sorts of companies worry about and do. In each class Bowen led the students in case discussions whose purpose was to understand the phenomena thoroughly. After a few classes, Bowen paused, and orchestrated a discussion through which they sought to define patterns in the phenomena – to begin categorizing by type of company, type of manager, and type of problem. Finally, they explored the association between these types, and the outcomes of interest. In other words, Bowen's course had an *inductive* architecture that moved up the theory pyramid. Then armed with their preliminary body of theory, Bowen and his students cycled down the deductive side of the pyramid to examine more companies in a broader range of circumstances. This allowed them to discover things that their initial theory could not explain; and to improve their constructs, refine their classification scheme, and improve their understanding of what causes what, and why.

There is another circumstance – where well-researched theories pertaining to a field of management already exist. In this situation, a *deductive* course architecture can work effectively. For example, Clayton Christensen's case-based course, Building a Sustainable Enterprise, is designed deductively. For each class, students read a paper that summarizes a normative theory about a dimension of a general manager's job. The students also study a case about a company. They then look through the lenses of the theory, to see if it accurately explains what historically happened in the company. They also use the theory to discuss what management actions will and will not lead to the desired outcomes, given the situation the company is in. Because the cases are complicated, students often discover an anomaly that then enables the class to revisit the categorization scheme and the associated statement of causality. Students follow this process, theory after theory, class after class, for the semester – and in the process, learn not just how to use theory, but how to improve it.[5]

---

processes that geneate numerical data the subjectivity that was exercised in the process of recording or not recording lies hidden.

5 At one point Christensen attempted to teach his course through an inductive architecture. Case by case, he attempted to lead his students to discover well-documented theories that prior scholars already had discovered. The course was a disaster – the wrong architecture for the circumstance. Students could tell that Christensen already had

As the experiences of Professors Bowen and Christensen suggest, the dichotomy that many see between teaching and research need not create conflict. It may be better to view developing and teaching courses as *course research*. And there are two circumstances in which professors might find themselves. When a body of theory has not yet coalesced, an inductive architecture is productive. When useful theory already has emerged, then a deductive architecture can make sense. In both circumstances, however, instructors whose interest is to build theory and help students learn how to use theory, can harness the brainpower of their students by leading them through cycles up and down the theory-building pyramid.

Implications: Theory as Method

Building theory in management research is how we define and measure our value and usefulness as a research community to society. We have focused on specific examples from management research to illustrate how our approaches to the empirical world shape what we can represent and can value and, more broadly, how theory collectively shapes the field of management research. This reminds us that building theory at an individual or collective level, handing off or picking up the baton, is not a detached or neutral process, yet the model developed here gives us a method to guide these efforts. From this model we recognize first the importance of both the inductive and deductive sides of the pyramid; second how subsequent cycles move us from attributes and substantive categories toward a circumstance-based understanding and more formal theory; and third eventually to an understanding of the relational properties that are of consequence and define the boundary conditions wherein the theory is of value.

This is our uiltimate aim: As students of business we readily accept that if employees in manufacturing and service companies follow robust processes they can predictably produce outputs of quality and value. When reliable processes are followed, success and failure in producing the desired result become less dependent upon the capabilities of individual employees, because they are embedded in the process. We assert that the same can be true for management researchers. If we follow a robust, reliable process, even the most "average" of us can produce and publish research that is of high value to academics and practitioners.

---

the answer, and his attempts to orchestrate a case discussion seemed like the professor was asking the students to guess what was on his mind. The next year, Christensen revised his course to the deductive architecture described above, and students reacted very positively to the same material.

Parking Lot for Important ideas that need to go somewhere:

So a major question that arises in conducting research is how do we know we are categorizing or measuring the best things to help us understand the phenomena of interest? Glaser and Strauss state that the elements of theory are, first, the conceptual categories with their conceptual properties and, second, the generalized relations among categories and their properties (1967: 35-43). A way to proceed with combining these elements is to emphasize a "relational" approach to theorizing (Bourdieu and Wacquant, 1992: 224-233) rather than just a substantialist approach. As already alluded to, a substantialist approach emphasizes "things" to be counted and categorized such as people, groups, products, or organizations. A relational approach, however, emphasizes the properties between things in a given area of interest, or what determines the relative positions of force or power between people, groups or organizations. The reason that most research follows a substantialist approach is that most methodological tools are focused on and best suited in identifying convenient sources data that can be easily counted and categorized more readily than the relational properties that exist between individuals, groups or organizations in a given social space over time (Bourdieu, 1989).

Given the methodological focus toward convenient sources of data to collect, it is not surprising that a substantialist approach dominates most of management research, as well as the social sciences. For example, the concept of "core competency" (Selznick, 1957) was developed to account for organizations that were successful in their environments. This concept became a very useful concept in the field of strategy in the late 1980s and the 90s (Prahalad and Hamel, 1990). However, the limitation of this category is that it was used to identify only successful companies; less successful companies were seen as lacking a core competency. The field of strategy did not begin to look more closely at the concept until Dorothy Leonard's research (1992; 1995) focus on the processes and outcomes that identified how a core competency can turn into a source of core rigidity. Leonard found that changes in a firm's "relations" to its suppliers and customers determine whether the firm can remain competitive. The corollary of this is that a core competency can become a core rigidity, diminishing competitive strength. By identifying this consequential "relations" Leonard not only provided a deeper formalization of "competency," but this also proved helpful to managers in suggesting how they apply their firm's resources to avoid this competency-rigidity tendency.

While a relational approach can push research to a deeper level of formalization, it raises methodological challenges. Because relations among individuals, groups or organizations are most telling as they change over time, a relational approach requires both the means of collecting data over time and a method of analyzing and representing the insights that such data can reveal. In one of the most influential ethnographic studies of technology implementation in management research, Barley's careful ethnographic analysis (1986; 1988; 1990) provided a comparative and temporal window into the implementation of the same technology in similar hospital settings. Despite these similarities, Barley documented very different outcomes in how radiologist and technicians joint used the CT-Scanning technology implemented. Based on these different outcomes, he asserted that technological and social structures mutually adapted differently over time. Barley observations over time helped to replaced the either or debate between the static view of technological determinism and the situated view of technology.

Using Barley's empirical documentation, Black, Carlile and Repenning (2003) formalized his observation at a more specific causal level through the use of a system dynamics method. This allowed them to specify the relation between radiologist and technicians and how their relative expertise in using the technology explains the different outcomes that Barley documented. Even though Barley recognized the importance of the "distribution of expertise" (Barley, 1986) between the two groups, he lacked a methodology to represent how over time the relative accumulations of expertise accounted for the different outcomes he observed. With this more formalized approach Black et al. could state a balance in "relative expertise" in using the new technology was essential in developing collaboration around a new technology. The specification of these relational properties was an improvement upon Barley managerial suggestion that a more decentralized organization is better able to successfully implement a new technology than a centralized one. This more formalized theory and relational understanding provides specific guidance to a practitioner about what to do when faced with the challenge of implementing a new technology when collaboration is desired.

This relational approach goes farther than a "contingency theory" approach (Lawrence and Lorsch, 1967)—because it recognizes not only are things contingent, but that in any situation some things, some relations, matter more than others in explaining the contingent (different) outcomes possible. The development of contingency theory has provided significant insight into the field of organizational behavior and design because it has identified that circumstances do affect outcomes. However, the fact that contingency theory is viewed by many as a stand-alone theory rather than a further reason to search for the particular sources of contingency limits the theory-building effort. This points to the proclivity of many researchers to leap directly from phenomena to theory and back again. If we continue around the theory building cycle, what we at first call contingent (e.g., decentralization versus centralization), upon further analysis reveals the underlying relational properties and why those relations are most consequential and why (e.g., how and why relative expertise matters).

## References


Allison, G. (197), *The Essence of Decision*. Glenview, IL: Scott, Foresman & Co.

Argyris, C. (1993), *On Organizational Learning*. Cambridge, MA: Blackwell.

Argyris, C. & Schon, D. (1976), *Theory in Practice*. San Francisco: Jossey-Bass.

Baldwin, C. and Clark, K.B. (2000), *Design Rules: The Power of Modularity*. Cambridge, MA: MIT Press.

Barley, S.R. (1986), "Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments." *Administrative Science Quarterly*, 31, 1: 78-108.

Black, L., Repenning, N. and Carlile, P.R. (2002) "Formalizing theoretical insights from ethnographic evidence: Revisiting Barley's study of CT-Scanning implementations." Under revision, *Administrative Science Quarterly.*

Bourdieu, P. (1989/1998), *Practical Reason*. Stanford: Stanford University Press.

Bourdieu, P. and Wacquant, L. (1992), *An Invitation to Reflexive Sociology*. Chicago: University of Chicago Press.

Bower, Joseph (1970), *Managing the Resource Allocation Process*. Englewood Cliffs, NJ: Irwin.

Bower, J.L., and Gilbert, C.G., eds. (2005), *From Resource Allocation to Strategy*. Oxford University Press.

Burgelman, Robert & Leonard Sayles (1986), *Inside Corporate Innovation*. New York: The Free Press.

Burgelman, Robert (2002), *Strategy Is Destiny*. New York: The Free Press.

Campbell, D.T.and Stanley, J.C. (1963), *Experimental and Quasi-experimental Design for Research*. Boston: Hougthon Mifflin Press.

Carlile, P.R. (2003), "Transfer, translation and transformation: Integrating approach in sharing and assessing knowledge across boundaries." Under revision, *Organization Science*.

Carr, E.H. (1961), *What Is History*? New York: Vintage Books.

Chandler, A. D. Jr. (1977), *The Visible Hand: The Managerial Revolution in American Business*. Cambridge, MA: Belknap Press.

Chandler, A. D. Jr. (1990), *Scale and Scope: The Dynamics of Industrial Capitalism*. Cambridge, MA: The Belknap Press.

Christensen, C.M. (1997), *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston: Harvard Business School Press.

Chesbrough, H.W. (1999). "The Differing organizational impact of technological change: A comparative theory of institutional factors." *Industrial and Corporate Change*, 8: 447-485.

Clegg, Brian (2003), *The First Scientist: A Life of Roger Bacon*. New York: Carroll & Graf Publishers.

Daneels, Erwin (2005), "The Effects of Disruptive Technology on Firms and Industries," *Journal of Product Innovation Management* (forthcoming special issue that focuses on this body of theory).

Gilbert, C.G. (2001), *A Dilemma in Response: Examining the Newspaper Industry's Response to the Internet*. Unpublished DBA thesis, Harvard Business School.

Gilbert, C.G., and Christensen, C.M. (2005). "Anomaly Seeking Research: Thirty Years of Development in Resource Allocation Theory." In Bower, J.L., and Gilbert, C.G., eds. *From Resource Allocation to Strategy*. Oxford University Press, forthcoming.

Fleming, L. and Sorensen, O. (2001), "Technology as a complex adaptive system: Evidence from patent data." *Research Policy,* 30: 1019-1039.

Glaser, B. & Straus, A. (1967), *The Discovery of Grounded Theory: Strategies of Qualitative Research*. London: Wiedenfeld and Nicholson.

Hayes, R. (1985), "Strategic Planning: Forward in Reverse?" *Harvard Business Review*, November-December: 111-119.

Hayes, R. (2002), "The History of Technology and Operations Research," Harvard Business School Working paper.

Hayes, R. and Abernathy, W. (1980), "Managing our Way to Economic Decline." *Harvard Business Review*, July-August: 7-77.

Hayes, R. and Wheelwright, S.C. (1984), *Restoring our Competitive Edge*. New York: John Wiley & Sons.

Hayes, R., Wheelwright, S. and Clark, K. (1988), *Dynamic Manufacturing*. New York: The Free Press.

Henderson, R.M. & Clark, K.B. (1990), "Architectural Innovation: The Reconfiguration of Existing Systems and the Failure of Established Firms." *Administrative Science Quarterly*, 35: 9-30.

Hunter, S.D. (2002), "Information Technology, Organizational Learning and Firm Performance." MIT/Sloan Working Paper.

Hutton, A., Miller, G., and Skinner, D. (2000), "Effective Voluntary Disclosure." Harvard Business School working paper.

James, W. (1907), *Pragmatism*. New York: The American Library.

Johnson, H.T. & Kaplan, R. (1987), *Relevance Lost*. Boston: Harvard Business School Press.

Kaplan, A. (1964), *The Conduct of Inquiry: Methodology for Behavioral Research*. Scranton, PA: Chandler.

Kaplan, R. (1986), "The role for Empirical Research in Management Accounting." *Accounting, Organizations and Society*, 4: 429-452.

Kuhn, T. (1962), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Lawrence, P. R. and Lorsch, J.W. (1967), *Organization and Environment*. Boston: Harvard Business School Press.

Leonard, D. (1995), *Wellsprings of Knowledge*. Boston: Harvard Business School Press.

Poole, M. & Van de Ven, A. (1989), "Using Paradox to Build Management and Organization Theories." *Academy of Management Review* 14: 562-578.

Popper, K. (1959), *The Logic of Scientific Discovery*. New York: Basic Books.

Porter, M. (1980), *Competitive Strategy*. New York: The Free Press.

Porter, M. (1985), *Competitive Advantage*. New York: The Free Press.

Porter, M. (1991), *The Competitive Advantage of Nations*. New York: The Free Press.

Raman, Ananth, (need citation)

Roethlisberger, F. (1977), *The Elusive Phenomena*. Boston: Harvard Business School Press.

Rumelt, Richard P. (1974), *Strategy, Structure and Economic Performance*. Cambridge, MA: Harvard University Press.

Selznick, P. (1957), *Leadership in Administration: A Sociological Interpretation*. Berkeley: University of California Press.

Simon, H. (1976), *Administrative Behavior* (3rd edition). New York: The Free Press.

Solow, R. M. (1985), "Economic History and Economics." *The American Economic Review*, 75: 328-331.

Sorensen, O. and Sorensen, J. (2001), Research Note - Finding the right mix: Franchising, organizational learning, and chain performance. *Strategic Management Journal*, 22: 713-724.

Sorensen, J. (2002), "The Strength of Corporate Culture and the Reliability of Firm Performance," *Administrative Science Quarterly*, 47: 70-91.

Spear, S.C. and Bowen, H.K. (1999), "Decoding the DNA of the Toyota production system." *Harvard Business Review*, September-October.

Stinchcombe, Arthur L. (1968), *Constructing Social Theories*." New York: Harcourt, Brace & World.

Sull, D. N. (2000), "Industrial Clusters and Organizational Inertia: An Institutional Perspective." Harvard Business School working paper.

Van de Ven, A. (2000), "Professional Science for a Professional School." In Beer, M. and Nohria, N. (Eds), *Breaking the Code of Change*. Boston: Harvard Business School Press.

Weick, K. (1989), "Theory Construction as Disciplined Imagination," *Academy of Management Review*, 14: 516-532.

Womack, J. P., Jones, D. T. & Roos, D. (1990), *The Machine that Changed the World*. New York: Rawson Associates.

Yin, R. (1984), *Case Study Research*. Beverly Hills: Sage Publications.

Yoffie, David, Sasha Mattu & Ramon Casadesus-Masanell (2002), "Intel Corporation, 1968-2003," Harvard Business School case #9-703-427.