Journal of
Organization Design

Vol. 1, No. 2 (2012)

ISSN 2245-408X

www.jorgdesign.net

TABLE OF CONTENTS

RESEARCH ARTICLES

TRANSLATIONAL ARTICLE

POINT OF VIEW ARTICLE

URGENT ISSUE

Journal of Organization Design
ISSN: 2245-408X

THE EVOLUTION OF ENTERPRISE ORGANIZATION DESIGNS

JAY R. GALBRAITH

Abstract: This article extends Alfred Chandler's seminal ideas about organizational strategy and structure, and it predicts the next stage of organizational evolution. Chandler described the evolution of vertical integration and diversification strategies for which the functional and multidivisional structures are appropriate. He also explained how the dominant structure at any point in time is a *concatenation*, or accumulation, of all previous strategies and structures. I extend Chandler's ideas by describing how early "structures" became "organizations" (people, rewards, management processes, etc.) and by discussing the more recent strategies of international expansion and customer focus. International expansion leads to organizations of three dimensions: functions, business units, and countries. Customer-focused strategies lead to four-dimensional organizations currently found in global firms such as IBM, Nike, and Procter & Gamble. I argue that the next major dimension along which organizations will evolve is emerging in firms which are experimenting with the use of "Big Data."

Keywords: Organization design; organization structure; strategy and structure

A major stream of thought in organization design is the evolution of the structure of the total enterprise. The origin of this idea is the historical study *Strategy and Structure* by Alfred Chandler (1962). This work, by a business historian, was picked up by organization theorists, strategic management theorists, economists, and sociologists. It led to a virtual explosion of conceptual and empirical studies of American, British, German, French, Italian, and Japanese enterprises (Franko, 1976; Stopford & Wells, 1972). Then, like many thought streams in organization and management theory, interest in it declined. In this article, I return to Chandler's concept of structural evolution and extend it to include today's global enterprise designs. This extension is based largely on my work as a practitioner, helping global companies as they develop the next phase of their growth strategies and enterprise structures.

STRATEGY, STRUCTURE, AND CONCATENATION

Chandler's idea that "structure follows strategy" is one of the best-known organizational concepts in business. His concept of concatenation, or accumulation, is virtually unknown. "The thesis... is then that structure follows strategy and that the most complex type of structure is the result of the concatenation of several basic strategies" (Chandler, 1962: 14). Concatenation drives the complexity of today's organizations but is also a management contradiction. Almost every leader is a champion of simplicity. But while leaders are saying "Keep it simple," they are acting to implement ever more complex strategies and structures. What is driving this contradiction?

Concatenation

Concatenation, the accumulation of simple strategies into increasingly complex structures, is at the core of Chandler's argument about structural evolution. Chandler explains the concept

Journal of Organization Design
JOD, 1(2): 1-13 (2012)
DOI: 10.7146/jod.6342

with the example of a start-up firm in a single location and with a single business function – such as a distributor. The first simple strategy emphasized by the distributor is volume expansion, which leads to a simple structure where an administrative office is created to manage the business. The next simple strategy is geographic dispersion. This new strategic emphasis results in adding a distribution department headquarters to administer the several distribution field units. The resulting, more complex structure is shown in Figure 1.

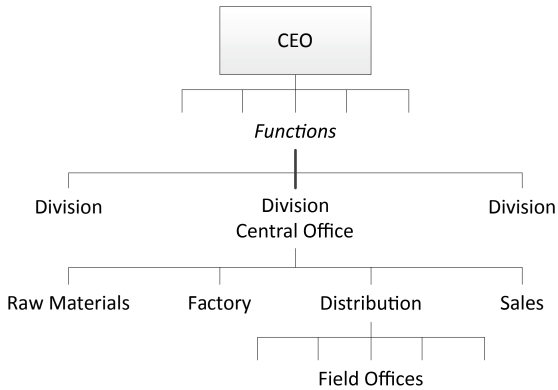

Fig. 1. Multidivisional Structure of Two Dimensions: Divisions and Functions

The next strategy is vertical integration, which adds additional functions to create a single, end-to-end business. The matching structure is a central office to coordinate the flow of work and material through the functions. This structure is also called a unitary form (U-form) or a multifunctional, single business structure (Williamson, 1975). The single business forms the basis of the next structure – the multidivisional or M-form. The multidivisional structure is created when the firm executes a strategy of diversification. The structure in Figure 1 shows the concatenation, or accumulation, of all the strategies in the more complex structure. In this case, each level represents a "simple" strategy, but the emphasis placed on a new simple strategy requires the development of a more complex structure. Table 1 shows the four strategies that Chandler discussed in *Strategy and Structure* (volume expansion, geographic dispersion, vertical integration, and diversification). The last two strategies, international growth and customer focus, were developed after Chandler's study and will be described below.

Table 1. Matching Strategy With Structure

Strategy	Structure
Volume expansion	Administrative office
Geographic dispersion	Departmental headquarters
Vertical integration	Division central office
Diversification	Multidivisional general office
International growth	Three-dimensional structure
Customer focus	Front/back and four-dimensional structures

Growth Drivers

As posed in the question above, why do leaders who prefer simplicity develop complex strategies and structures? There appear to be two main drivers of this behavior. The first is the pursuit of growth. Every publicly traded company wants to grow and drive its stock to trade at a premium. A high stock price makes it easier to attract capital and reward executives and to serve as a currency to make acquisitions. Also, talented people want to join a growth company that has a bright future. But while growth is desirable, it is also limiting. A firm can only grow so much in its home country and core business. Therefore, it must venture into new country markets and adjacent businesses to attain growth. In so doing, the firm increases the complexity of its strategy and structure.

The other growth driver is the Law of Requisite Variety (Ashby, 1956). Taken from control theory or cybernetics, the law states that as entities in the stakeholder environment proliferate, units inside the enterprise must also proliferate to respond to those entities. For example, the marketing function has evolved from dealing with the mass market to focusing on market segments and now micro-segments. Large food companies (and their marketing consultants) may focus on more than 650 micro consumer segments. Some of these segments, such as Hispanic mothers, senior foodies, or the freezer segment, exhibit faster growth than the food market as a whole. Therefore, the food companies develop new products and promotions for those growing micro segments. Inside the marketing function, people and new departments are assigned and created to manage these chosen new micro segments. In this way, proliferating entities in the stakeholder environment lead to proliferating organizational units in response.

As long as companies pursue growth and stakeholder environments increase in complexity, enterprise organizational structure will evolve into more complex forms. Before describing the next strategies that have led to more complex organizations, let's trace the changes to the multidivisional form since Chandler's study.

EVOLUTION OF THE MULTIDIVISIONAL STRUCTURE

Chandler originally described and studied three types of structures: functional, multidivisional, and holding company. For each structure, there was a unique strategy. A functional organization was the means for implementing a single business strategy. A multidivisional structure was used for diversification into multiple related businesses. The holding company was appropriate when diversifying into multiple unrelated businesses. Researchers called these structures the unitary or U-form, multidivisional or M-form, and holding company or H-form (Williamson, 1975). Since Chandler's book was published in 1962, these forms have evolved from "structures" to complete "organizations" and from pure to mixed forms.

From Structure to Organization

The first additions to Chandler's basic strategy-structure relationship were the processes, policies, and practices used for managerial control and coordination. These additional features enlarged the organizational structure into a complete organizational form. Organization designers then wanted a model that would help them design and build organizations. In my case, I began to use the Star Model, shown in Figure 2, to guide my design thinking and practice (Galbraith, 1977, 2002). The Star Model has five key components for building an organization. McKinsey, a management consulting firm, created the 7S Model, which has seven organizational building blocks (Peters & Waterman, 1982). Today, every management consulting firm has a version of the Star Model or the 7S Model to guide its own design practice.

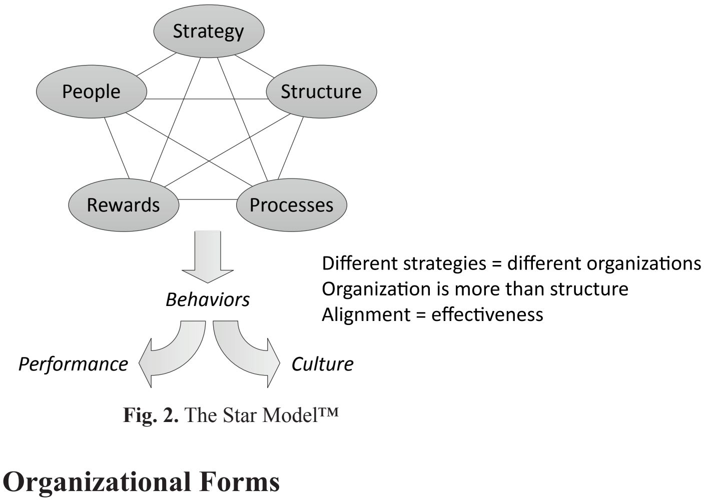

Fig. 2. The Star Model™

From Pure to Mixed Organizational Forms

When organization designers tried to design structures for diversified companies, they found it difficult to distinguish between related and unrelated portfolios. The M-form and H-form and their strategies were pure types, but many actual organizations were mixtures of both. Designers' thinking evolved from drawing the line between related and unrelated strategies to conceiving of portfolio strategy as a continuum from a single business to multi-business unrelated portfolios. The related multi-business portfolio was somewhere in between. Some attempts were made to measure the continuum. One measure was the two-digit level of Standard Industrial Classification (SIC) codes. However, SIC codes proved to be not very useful. A company like Proctor & Gamble measured high in diversification based on SIC codes because it operates in many product categories such as soap, paper, pharmaceuticals, and so on. Yet P&G was a classic M-form organization because every profit and loss (P&L) center was a B2C business in the packaged consumer goods sector. All the P&Ls at P&G followed the same business model. So today, researchers and designers make judgments about where a portfolio is located along the low-to-high portfolio diversity scale. These judgments are based on the number of similar businesses.

A new type of organization emerged because the gap between related and unrelated business portfolios was still very large. Thus, a mixed model was created from among the three pure types. For example, originally Hewlett-Packard had a classic related portfolio of P&Ls (which it called divisions). The company provided electrical instruments for technically trained professionals in the electronics, telecom, chemical, and medical industries. Then HP diversified into computers. At first it followed the same business model of selling minicomputers (boxes) to end users in factories and laboratories. However, the computer business eventually evolved into a "product" and "systems" business. Customers would buy an entire supply chain system made up of many computers, storage devices, software, and services, while computer products like PCs, printers, and hand-held devices were sold as stand-alone products through resellers. As a result, HP wound up with three different business models for (1) computer systems, (2) computer products sold through resellers to businesses and consumers, and (3) instrument businesses selling boxes directly to end users. Such business model diversity made HP look like some unrelated corporate portfolios. The "HP Way," on the other hand, made the company look like some related-portfolio companies. Thus, HP fell into a mixed category.

Today, nearly every consulting firm uses a four-category continuum to analyze and evaluate corporate portfolio strategies. For each of the four types of strategies, there is a complete Star Model including the appropriate structures, processes, and HR practices. The entire model is usually shown in a grid with the strategy types on one axis and the organizational elements, which match that strategy, shown on the other axis. The grid that I use is shown in Table 2 (Galbraith, 1993).

From Divisions to Strategic Business Units

As multidivisional corporations grew, they initially used a cell-division model to adjust their organizations. That is, when a division reached a size of about \$150 million in revenues, it was divided into two divisions of about \$75 million each. When these two divisions grew to be \$150 million, each of them was divided, thereby creating four divisions of \$75 million each. General Electric followed this process of cell division as it grew. In 1969, the result was that it had a three-level structure reporting to the CEO's office. The first level consisted of eight business groups. Reporting to the groups were 49 divisions, and reporting to the divisions were about 250 departments. The departments were the basic building blocks and P&L centers of the structure. When they attained revenue of approximately \$150 million, they were split into two departments. At that time, the CEO called in McKinsey consultants to help reduce the complexity of the organization. There was simply no way that the CEO and the leadership team at GE could understand the operations of 250 departments and allocate resources to them during the strategic planning and budgeting process.

Table 2. Portfolio Strategy and Organization

Strategy	Single Business	Related Diversification	Mixed	Unrelated Diversification
Structure	Functional	Divisional	Cluster	Holding Company
Centralization	High	Moderate	Low to the cluster Moderate within cluster	Low
Corporate Staff	Small	Large	Low in corporate Moderate in cluster	Small
Control Type	Operational Strategic Financial	Strategic Financial	Strategic Financial	Financial
Business Processes	Common	Common	Common within cluster	Different
Compensation System	Company	Company	Cluster	Subsidiary
Bonuses	Company	Company	Cluster	Subsidiary
Careers	Company	Company	Mix	Subsidiary
Subsidiary Culture	Company-wide	Company-wide	Mix	Unique to subsidiary
Division Name or Brand	Company-wide	Company-wide	Mix	Subsidiary
Example	BMW	Agilent	Hewlett-Packard	Berkshire Hathaway

The result was the GE-McKinsey SBU reorganization of 1969–1970. In this reorganization, the divisions and departments were collapsed into 49 strategic business units (SBUs), which reported to eight group executives who reported to the CEO's Office. An SBU was now the basic building block of the organization. The major change in thinking here was that size was no longer the determinant of the organizational building blocks. Those blocks were now determined by a business rather than size logic. An SBU was to be a complete, fully functional business with its own unique set of products, customers, and markets. SBUs varied in size from \$100 million to about \$1 billion in revenue. The SBU logic and language subsequently spread from GE to many of today's large corporate enterprises.

Another important outcome of the 1969–1970 reorganization was the dividing of the Personnel or Employee Relations Group into two parts, one for blue-collar workers and their unions and one for senior executives (the top 250 people). The new Executive Resources unit would focus on the development and compensation of these 250 senior executives. A talent review process was also added and implemented. The top 250 executives were seen as talent belonging to GE the corporation, not just to the SBUs where they worked. The talent pool contained executives who could be moved around the enterprise to where they were needed. Such resource flexibility is a major advantage when responding to changing market opportunities.

In summary, since the publication of Chandler's (1962) book, organization designers have evolved in their thinking along the following lines:

(a) from strategy and structure to strategy and organization
(b) from three pure types of strategy and structure – U-form, M-form, and H-form – to a continuum of strategies and organizations where the pure U-form is the simplest and the pure H-form is the most diverse

(c) from divisions, where size determines the basic building blocks of the organization, to SBUs, where business logic determines the basic building blocks of enterprise structure.

In the next section, we build on Chandler's work to see how international expansion led to the next concatenation of enterprise organization.

INTERNATIONAL EXPANSION AND THE DEVELOPMENT OF THE MULTIDIMENSIONAL MATRIX STRUCTURE

As noted above, firms driven by growth inevitably reach the limits of their core business in their home country. Chandler (1962) reported that American firms in the mid-1900s chose to grow by diversifying within the U.S. It was at this time that managers created and adopted the M-form and H-form organizations. But as U.S. domestic growth declined in the 1960s and the European Common Market opened up, American firms expanded into Canada and Europe and then into other regions of the world. Our knowledge of this phase of strategy and structure development comes largely from Raymond Vernon's Harvard International Project at the Harvard Business School. The research was reported by Stopford & Wells (1972), Franko (1976), Prahalad & Doz (1987), and Bartlett & Ghoshal (1989). At first, the growth strategy of international expansion caused few changes in the M-form companies that were the first to expand. Those companies simply added another division, the International

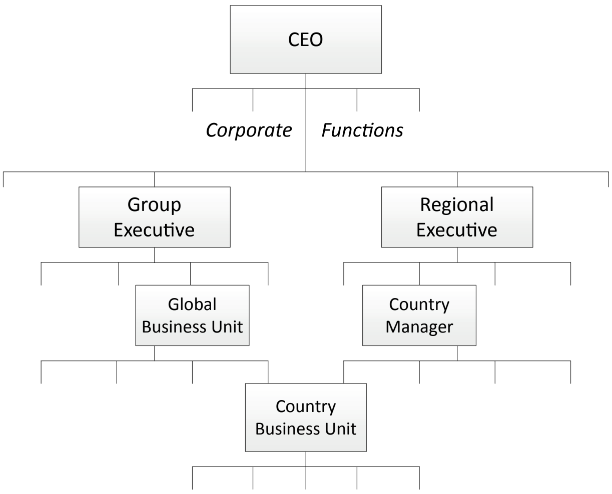

Fig. 3. Multi-Business, Multinational, Functional Matrix

Division, to their existing organization. But when the international division's sales reached 25-35% of total sales, it was disbanded and a new form of structure emerged (see Figure 3). This structure was a three-dimensional structure wherein the business units, countries, and functions all reported to the CEO. The M-form and H-forms were two-dimensional structures wherein the businesses were the P&Ls and reported to the CEO along with functions.

The variations in the three-dimensional structure concern the placement of the P&L responsibility as well as power and authority relationships. For companies with SBUs in B2B businesses and which spend more than four percent of sales on R&D, the P&Ls are in the business dimension, as in Royal Dutch Shell. For companies in B2C businesses and which spend much less on R&D, the P&Ls are in the geographic dimension, as in Nestlé. Still other firms, such as ABB, introduced a balanced matrix organization with the P&Ls in both the businesses and the countries. ABB was a B2B company that spent seven percent of sales on R&D. These features favored a business dominant structure. But 70% of ABB's sales went to government-owned or -influenced customers. Such customers required a strong local presence in their countries and a strong country manager. ABB had to be both globally integrated and locally responsive. A dual profit center matrix was its chosen structure.

The functions are the third dimension of a matrix structure. The functions of finance and R&D were the most global units and were matrixed across the P&Ls. As supply chains became more integrated across borders, the supply chain function gained more power and authority. Similarly, global brands increased the power of marketing. Balancing the power and authority of functions, countries, and businesses is the challenge for leaders and management teams in the three-dimensional matrix (Bartlett & Ghoshal, 1989; Galbraith, 2000).

In summary, a new strategy of international expansion led to a new three-dimensional organization. As in the past, the resulting organization was a concatenation of past strategies and structures. Initially, an international division was simply added on to the existing multidivisional structure, but increased growth led to embedding the geographical activities into the functions and businesses. The result was a complex three-dimensional organization, which is often some form of matrix structure. Getting these complex structures to work effectively is still an organizational design challenge for companies today.

THE CUSTOMER DIMENSION

In the 1990s, a new customer-focused strategy began to emerge, and it continued to be refined during the early years of the 21st century. There are a number of drivers behind this new strategy, but two are particularly salient. One is the shift of more buying power into the hands of customers. Big customers are demanding and getting a single interface with their vendors and are receiving customized product and service offerings to meet their particular needs. The other driver is the move to provide systems or solutions to customers instead of stand-alone products. The digital revolution now allows every object to talk with every other object. So, for example, IBM provides smart solutions to customers like a smart electrical grid for the island of Malta. The issue with providing unique offerings to customers is that vendors lose economies of scale. Their response has been mass customization. Now top companies design products and/or solutions platforms that can be replicated around the world (Gawer & Cusumano, 2002). These platforms have been designed from the very beginning to be easily and quickly modified to meet the unique needs of customers. The organization for

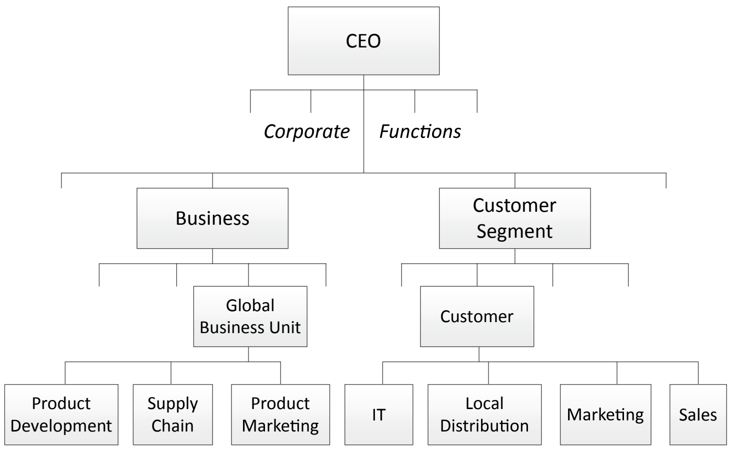

Fig. 4. Front/Back Organization

implementing a mass-customized product design strategy is the front-back model (Galbraith, 1993, 2002: Ch. 8).

The front-back organization is shown in Figure 4. It is a modification of the matrix design shown in Figure 3. The front end of the business is organized around customers, countries, and/or customer segments, while the back end is organized around products. The back end is to achieve global scale, while the front end achieves local adaptation. The supply chain, product marketing, and product development functions are more globally and regionally oriented in the front-back model. There is no country-business matrix. The customer-centric front end focuses on developing in-depth customer knowledge, cross-selling, and custom solutions. Marketing is local and focused on segmentation, local distribution, and customer insights. Either or both ends of the organization can be the P&L centers. At some companies, such as Unilever, the organization is three-dimensional like the one shown in Figure 4. In other firms, such as P&G, Nike, and IBM, the customer or customer segment is a fourth dimension.

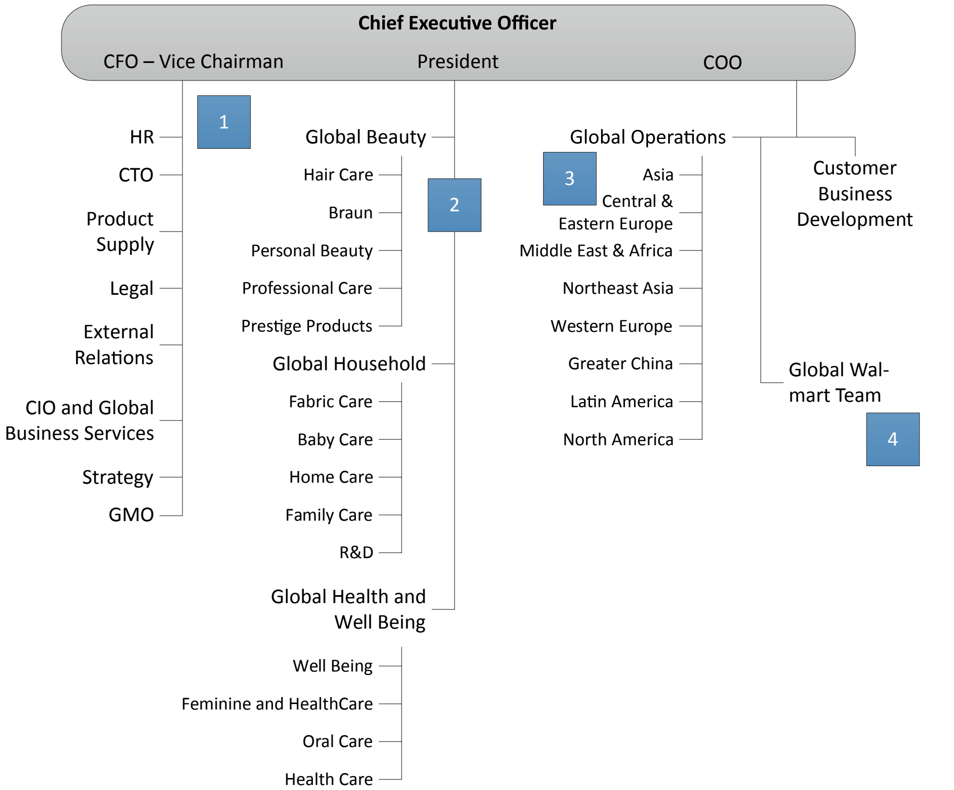

Fig. 5. P&G's Four-Pillar Organization

The Four Pillar organization of P&G is shown in Figure 5. Reporting to the Office of the Chief Executive are (1) business functions, (2) global business units, (3) market development organizations (regions), and (4) the Wal-Mart customer team (Wal-Mart is one of P&G's many large customers). Wal-Mart accounts for about 35% of P&G's worldwide sales, and the customer team consists of about 250 people located at various Wal-Mart sites around the world. In this organization, the Global Business Units (GBUs) design the global platforms, and the regions and customer teams adapt them for their customers. For example, the Hair Care GBU creates the platform for Pantene Shampoo. In Japan, part of Northeast Asia, the local product manager ensures that less perfume is added (Japanese women prefer more subtle scents) and more conditioner is added (Asians have thicker hair). Similar customizations are made for global customers such as Tesco and Carrefour. The GBU handles the mass aspects of a product, and the regions and customer teams are responsible for the custom portion.

Many companies have responded to the increase in customer power by creating global account units. Under pressure from Daimler-Benz, ABB created a global account team consisting of seven core members, those with the largest sales volume, and 30 extended team members. The team provided a single interface for Daimler to jointly create a global sales plan for the two companies to work together. Such global account teams are usually limited to the sales function.

The customer teams at P&G, Nike, and IBM are multifunctional units, which are often P&L-responsible as well. The customer dimension in these companies has become a fully embedded fourth dimension in their organizations. A partial example of embedding through the matrix relations at P&G is shown in Figure 6. The usual matrix of corporate functions across business units is shown only for the Chief Technical Officer (CTO) and one business group. The Wal-Mart team's matrix reporting is more elaborate. It shows how the customer dimension is embedded into the other three dimensions of the organization. The regional team reports to the North America region as well as the global Wal-Mart team. The business team in North America (NA) reports to the NA Fabric Care unit as well as the regional Wal-

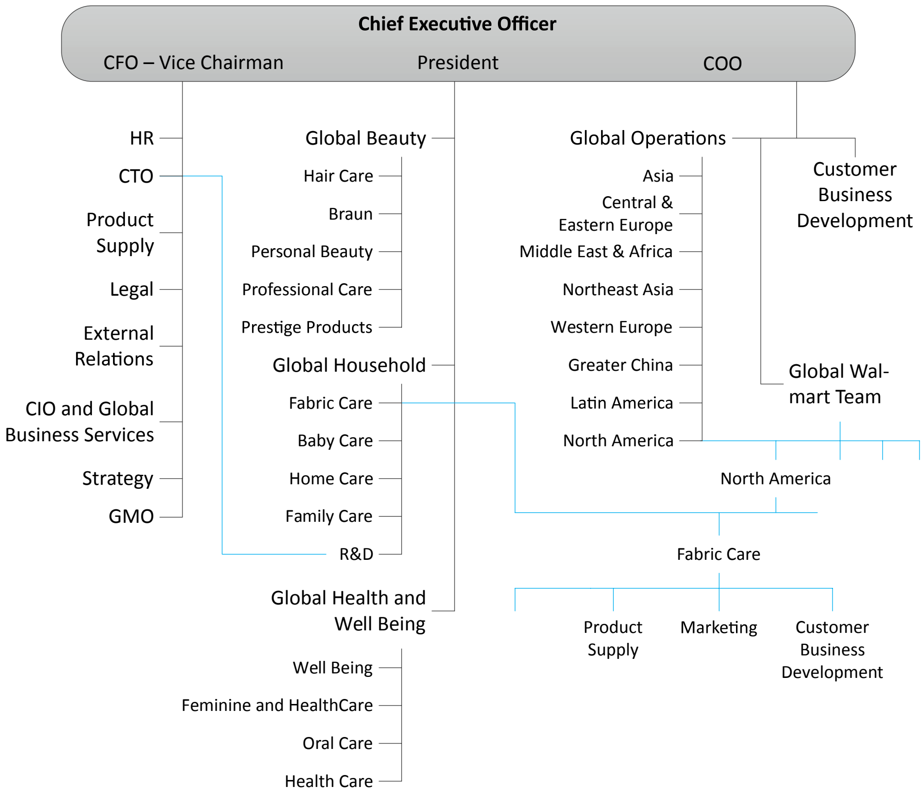

Fig. 6. Matrix Relationships at P&G

Mart team. Also, the functions report to the Wal-Mart NA Fabric Care unit. Even with this simplified diagram, the complexity of the four-dimensional organization is obvious. Thus, it is clear why organization designers began to see that organization is more than structure. How does P&G make this complex organization work? One major means is through aligning goals in the planning process. The regional Wal-Mart teams align their revenue, margin, and growth goals with the regions; the Wal-Mart regional business team aligns its goal with the GBUs; and so on. P&G also uses rotational assignments extensively. People join P&G in a function and then rotate through GBUs, regions, customer teams, and corporate headquarters. They build personal networks and get an understanding of the total company and how it works. They can then participate in the processes to align the company across the four dimensions. So as organizations get more complex, the more important are the management processes and HR practices. They provide the means for holding global enterprises together.

SUMMARY: EVOLUTION OF ENTERPRISE ORGANIZATION

I have tried to extend Chandler's original concept of strategy and structure to encompass current thinking and practice regarding enterprise-level organization. Chandler (1962) described how the strategies of vertical integration and diversification led to the single-business functional structure and the multi-business divisional structure. Following Chandlerian logic, I described how the multidivisional structure was expanded to *organization* by including management processes and HR practices. Following this, I described the GE reorganization of 1969–1970 (which was heavily influenced by the thinking of McKinsey consultants). This reorganization shifted thinking away from size as being the shaper of divisions to business logic being the shaper. The next strategy that affected enterprise organization was international expansion. The Harvard International Project provided the bulk of the research that led to our understanding of the resulting three-dimensional organization (functions, businesses, and countries). Finally, I described the most recent strategy of customer focus. This strategy has led to the four-dimensional organization where businesses, functions, countries, and customer units all report to the company's top leadership.

Throughout the discussion, I have tried to illustrate Chandler's little-known concept of *concatenation*. This concept suggests that at any point in time, a company's organization is the accumulation of all previous strategies and structures the company has adopted. Concatenation leads to organizational complexity because each new strategy is added to prior strategies, which have become institutionalized. To be sure, companies can shrink and divest and otherwise move in the opposite direction. For example, circa 1990, Westinghouse was a three-dimensional global company. Today it is a service provider to nuclear power plants under Japanese ownership. But for the IBMs and P&Gs of the world, where growth is paramount, concatenation is a reality that must be addressed by organization design theory and practice.

The P&G organization shown in Figure 6 is a result of concatenation. Up until World War II, P&G was in a single business (soap). That business is now the Family Care Global Business Unit. P&G then diversified into detergents (Fabric Care), paper (Home Care, Baby Care), shampoos (Hair Care), and so on. The GBUs are collected into three groups, as shown in Figure 6. In the 1960s, P&G expanded into Europe and subsequently the rest of the world. It added regions as it grew and linked them to the GBUs. In the late 1980s, P&G formed an alliance with Wal-Mart. It also built alliances during that period with other fast-growing customers in North America. As Wal-Mart grew internationally, so did P&G. Around 2000, P&G implemented its Organization 2005 initiative. This organization was the four-dimensional, front-back model shown in Figures 5 and 6, also referred to as the Four Pillar organization. Over time, P&G has embedded the customer teams into its GBUs and regions. Currently, P&G obtains well over half of its total sales from ten global retailers who are its customers.

FUTURE CONCATENATION

Given the complexity of global enterprise organizations, one must ask, Is there any end to this process? Will companies continue to add new strategic dimensions and embed them in their organizations? Using Chandlerian logic, there are actually two parts to this question. First, will the growth drivers continue to create additional organizational dimensions as markets, channels, and media become more fragmented and specialized? And, second, can organizations continue to create the integrating mechanisms needed to handle more complex interdependence?

The addition of strategic dimensions is a slow process. During the nineteenth century, Chandler (1962) identified three business growth strategies: volume expansion, geographic dispersion, and vertical integration. During the twentieth century, three more growth strategies were developed: diversification, international expansion, and customer focus. So, only six major growth strategies have been driving organizational evolution over the past one hundred fifty years. Assuming that the limit of strategic dimensions has not yet been reached, what is the next one likely to be? One candidate that is rapidly emerging is "Big Data."

Big Data

Big Data is a combination of multiple large databases within a firm, the continuing advance of cost-effective storage and computing hardware, and the use of analytics to make sense of all the data. Initially, data management was seen as an IT responsibility. The analytical expertise, however, usually resides in the businesses and functions. So today, the capability to effectively use Big Data is becoming an enterprise-wide responsibility. SAS, a software vendor that sells business intelligence software packages, is promoting the establishment of BICCs, business intelligence competency centers. BICCs are permanent structures that are staffed from across the company and make the data and analytics universally available. Some companies are now emerging as leading Big Data competitors. For example, Capital One has always been an information-based company. The firm started with a credit card business and has diversified into financial services of all types. Its data and analytics competence is used in all of the firm's business units. American Express talks about moving from a "payments" company to a full-fledged information company, with possibly many new businesses emerging if it succeeds in being able to capture and analyze huge amounts of information.

There are a number of signs indicating that Big Data may be the next strategic dimension and a new source of growth. McKinsey Global Institute (2011) published an important report entitled “Big Data: The Next Frontier for Innovation, Competition, and Productivity.” McKinsey has a practice area in Big Data, and several articles have appeared in the McKinsey Quarterly. The World Economic Forum held a session on Big Data at its 2012 Davos meeting and issued a report entitled “Big Data, Big Impact: New Possibilities for International Development.”[1] IBM, like SAS, has focused its Smart Planet initiative on helping companies become better competitors by using Big Data strategically.

The evolution of Big Data as a strategic dimension has important organization design implications. In their book, *Competing on Analytics*, Davenport and Harris (2007) identify five stages through which an ordinary company must pass in order to become an analytical competitor or master of Big Data. In the earliest phase, there is the creation of the enterprise-wide database. Initially, different units all have their own databases. Finance has credit and risk data, manufacturing and procurement have vendor data, and marketing and sales have customer and channel data. All of these databases need to be combined into one enterprise data network with access for all. Then there is the maneuvering of various units to become the central data and analytics unit. At a bank, for example, there are several candidates for this role: the Chief Risk Officer who has the credit data and risk analytics, the Chief Marketing Officer who has the customer data and customer insight analytics, and the credit card division that has credit, marketing, and merchant data. While they are all competing to be the central unit, they are all united around opposing a central unit reporting to the CEO. All such companies will need organization design and change expertise to develop acceptable solutions.

One real (anonymous) company shows how a firm can adapt organizationally to the potential opportunities afforded by Big Data. It started as a credit bureau with access to all bank credit transactions and developed the analytical capabilities to calculate credit scores. It grew by acquisition, acquiring both other credit bureaus and marketing database companies. Today, the company can combine these business databases and sell to retailers lists of consumers who have both the financial ability and the willingness to buy. Early on, this company created a Data Council to merge or network all of the various databases of its acquired companies. It also bought or acquired outside data. For example, all automobile transactions are recorded by state Departments of Motor Vehicles, and these data are publicly available. The company has grown to consist of multiple credit and marketing businesses in over 25 countries. It provides data services through multiple customer channels to banks, credit card companies, insurance companies, telecom operators, and retailers. Anyone who grants credit to consumers is a potential customer. It is the type of four-dimensional organization (functions, businesses, countries, and customers) described earlier as in an advanced stage of structural evolution. Now there is a fifth dimension, represented in an organizational unit called Decision Analytics which provides services for all of the firm’s businesses and countries. Decision Analytics develops fraud analytics to distinguish late payers who are down on their luck from those who want to commit theft. It has predictive analytics that can tell when a consumer is likely to switch to a new credit card company. This same product is now valuable to a telecom operator. Wireless telecoms are plagued with “churn.” Since telephone numbers were made portable, consumers continually move from one telecom operator to another. The company’s software can inform operators in advance of churn so that they might offer customers something that will retain them. This company is becoming a user, and ultimately perhaps a master, of Big Data. It has integrated all of its data in a company-wide network and is developing real skills in combining databases and analyzing them to create unique customer insights. Big Data has allowed it to grow during the downturn.

In summary, it seems to me that the concatenation process will continue. My leading candidate for the next growth driver and strategic dimension is Big Data, the use of massive databases and analytics on an enterprise-wide basis. The Big Data phenomenon has been emerging over the past 10-15 years and appears to be gathering momentum.

---

1 The report can be downloaded at: http://www.weforum.org/reports/big-data-big-impact-new-possibilities-international-development

Integrating Big Data into the Organization

A consequence of the concatenation of strategic dimensions is the ever-increasing need for more integration. Can organization designers create the integrating mechanisms that can coordinate the increasing interdependence of today's complex organizations? With the era of Big Data upon us, it may be that Big Data itself will create the tools necessary to manage interdependence.

Many of the mechanisms to manage interdependence have been both formal processes and automated processes. For example, Cisco Systems customers can go to Cisco's website and design their own order for routers. At a click of the mouse, the order is sent into Cisco's supply chain system. The customer's credit is checked, and orders are sent to contract vendors who manufacture parts and send them to the contract assembler. When completed, the order is shipped and an invoice is sent to the customer. Upon receipt and validation of the order, the customer wires the money to Cisco's bank account. The vendors, shippers, and taxes are paid by the same process. In such a system, people only manufacture and assemble the product. In earlier days, sales people would have entered the order, and finance and accounting people would have checked the credit, recorded the order, and sent it to production scheduling. People would have scheduled the orders, talked with vendors, and then sent the necessary paperwork. Today, there are very few people involved in this process. Fewer people mean fewer managers, fewer departments, and no face-to-face coordination. Interdependence is managed through automated processes. Coordination and integration processes are now so extensive that they have been called the Second Economy (Arthur, 2011).

Second, there have been advances on the management front. For example, resource allocations are being made using spreadsheets that show businesses as the rows and countries as the columns. The budgeting process is designed to arrive at shared goals for both a business and a country. These spreadsheets then become the "dashboards" to manage changes throughout the year. There are times when a large meeting is employed to create and update the spreadsheets. Using the concept of "get the system in a room," groups of 27, 43, or 71 people are gathered into "decision accelerators" and facilitated to reach a decision. Interdependence is managed through such emerging integrating mechanisms called horizontal processes (Galbraith, 2010).

Another management area where integration has advanced is HR systems. Increasing integration requires more people who are able to work together collaboratively. These kinds of people are recruited and promoted to key decision-making positions. The use of 360-degree evaluations and coaching has helped people develop their collaborative skills. Rotational assignments have increased in importance. Many HR practices today emphasize enterprise-wide collaboration. When people have rotated through various parts of the company, they develop a comprehension of the entire company and how it works. At P&G, one enters the firm through a function. People stay in that function but work periodically at the functional headquarters, in a business unit, in a country, and on a customer team. Leaders who complete this process identify with the company and know its culture and systems.

Finally, shared values guide people to choose behaviors that result in desired outcomes without having to converse and coordinate with others. Think how much more effective the mortgage origination firms would have been if they had a value of "Never sell a mortgage to a person who cannot afford one." With reinforcement from management, values serve as self-monitoring mechanisms to guide people's behavior in desirable ways.

In summary, the automation of processes will coordinate a lot of interdependent work. The development and selection of collaborative leaders who can participate in large-scale meetings will manage the interdependence in the resource allocation processes. And the continued use of rotational assignments will develop leaders with a good network and a "one company" mindset. Combined, these practices will enable companies to manage their increasing interdependence and complexity. Companies that cannot develop these capabilities and practices will miss out on the next source of growth.

CONCLUSION

The evolution of complex enterprise organization continues. From the single-business functional structure, we have evolved to the four-dimensional organization adopted by many leading firms. The natural question is "Will there be a fifth dimension?" From early indications, the answer appears to be "Yes." A number of organizations are integrating previously isolated databases into an enterprise-wide database. Then, using new analytical tools, they are discovering useful insights that have value to customers. The move to Big Data appears to be the next dimension in the concatenation process that Chandler identified. Big Data will then become the next big challenge for organization design.

REFERENCES

Arthur WB. 2011. The second economy. *McKinsey Quarterly 4*.

Ashby WR. 1956. *An Introduction to Cybernetics*. Wiley, New York.

Bartlett C, Ghoshal S. 1989. *Managing Across Borders*. Harvard Business School Press, Boston.

Chandler AD. 1962. *Strategy and Structure*. MIT Press, Cambridge, MA.

Davenport TH, Harris J. 2007. *Competing on Analytics*. Harvard Business School Press, Boston.

Franko L. 1976. *The European Multinationals*. Greylock Press, Greenwich, CT.

Galbraith J. 1977. *Organization Design*. Addison-Wesley, Reading, MA.

Galbraith J. 1993. The value adding corporation. In J. Galbraith and E. Lawler (Eds.), *Organizing for the Future*: 15-42. Jossey-Bass, San Francisco.

Galbraith J. 2000. *Designing the Global Corporation*. Jossey-Bass, San Francisco.

Galbraith J. 2002. *Designing Organizations*. Revised Edition. Jossey-Bass, San Francisco.

Galbraith J. 2010. The multi-dimensional and reconfigurable organization. *Organizational Dynamics* 39: 115–125.

Gawer A, Cusumano MA. 2002. *Platform Leadership: How Intel, Microsoft and Cisco Drive Industry Innovation*. Harvard Business School Press, Boston.

McKinsey Global Institute. 2011. Big Data: the next frontier for innovation, competition, and productivity.

Peters TJ, Waterman RH Jr. 1982. *In Search of Excellence: Lessons from America's Best-Run Companies*. Harper & Row, New York.

Prahalad CK, Doz Y. 1987. *The Multinational Mission*. Pergamon Press, New York.

Stopford J, Wells C. 1972. *Managing Multinational Enterprise*. Longmans, London, UK.

Williamson OE. 1975. *Markets and Hierarchies*. McGraw-Hill, New York.

JAY R. GALBRAITH

Professor Emeritus
IMD, Lausanne, Switzerland
E-mail: jay@jaygalbraith.com

THE VIRTUAL DESIGN TEAM

DESIGNING PROJECT ORGANIZATIONS AS ENGINEERS DESIGN BRIDGES

RAYMOND E. LEVITT

Abstract: This paper reports on a 20-year program of research intended to advance the theory and practice of organization design for projects from its current status as an art practiced by a handful of consultants worldwide, based on their intuition and tacit knowledge, to: (1) an "organizational engineering" craft, practiced by a new generation of organizational designers; and (2) an attractive and complementary platform for new modes of "virtual synthetic organization theory research." The paper begins with a real-life scenario that provided the motivation for developing the Virtual Design Team[1] (VDT), an agent-based project organizational simulation tool to help managers design the work processes and organization of project teams engaged in large, semi-routine but complex and fast-paced projects. The paper sets out the underlying philosophy, representation, reasoning, and validation of VDT, and it concludes with suggestions for future research on computational modeling for organization design to extend the frontiers of organizational micro-contingency theory and expand the range of applicability and usefulness of design tools for project organizations and supply-chain networks based on this theory.

Keywords: Virtual design team; project organization design; organization design

MOTIVATION FOR PROJECT ORGANIZATION DESIGN THEORY, METHODS, AND TOOLS

In 1987, Art Smith, the vice president in charge of facilities for a major semiconductor manufacturer, "Micro," was facing a significant organization diagnosis and design challenge. The product life cycle of a new microprocessor is very short – three to six months – before either a competitor or Micro itself produces an even faster microprocessor, at which time the price of that generation of microprocessors must be discounted, so that its gross margin falls significantly from its original level of around 60%. Each production train for a new microprocessor was producing about $1 million of product per hour for Micro early in its life cycle at that time, and a typical fabrication facility (fab) contained three production lines. Any delay in completing a fab on its planned date would cost Micro about 60% of three million dollars per hour of gross margin, seven days per week, 24 hours per day. Thus, on-time completion of a fab was an exceptionally high priority for Micro.

Exacerbating Art Smith's challenge, Micro's manufacturing engineers insisted on waiting until the last possible moment to order the rapidly evolving manufacturing equipment for its fabs, in order to avoid having obsolete equipment in the fabs from day one. Each piece of manufacturing equipment in a fab has different requirements with respect to the geometric layout for moving the silicon wafers between machines, its mounting geometry, the structural support it requires, the fluids and gases to be supplied to it, etc. The detailed design and

1 The Virtual Design Team (VDT) research described in this paper has been supported at different times by the Center for Integrated Facility Engineering and Collaboratory for Research on Global Projects at Stanford University, the National Science Foundation, and the Center for Edge Power of the Naval Postgraduate School. The support of these organizations for the VDT research is gratefully acknowledged. However, the author is solely responsible for the opinions expressed in this paper

DOI: 10.7146/jod.6345

construction of the fab must proceed extremely rapidly and concurrently once the specific new equipment has finally been selected. At the same time, the date at which the fab needs to begin producing microprocessors in quantity is planned far in advance to match the time at which the semiconductor design will be finalized, the photolithography masks for etching the chips will be ready, and the marketing plan will be in place, so that the microprocessors can hit the market in large volume and with high quality at just the right moment.

As Micro's manufacturing engineers pressed Art's team to delay equipment purchases ever closer to the fixed fab completion dates, the fab design and construction projects came under extreme schedule pressure. Micro's response to this pressure was to schedule many highly interdependent tasks concurrently. As the tasks were executed more and more concurrently, the fab delivery projects began to experience an exponentially larger volume of design changes and rework, resulting in delays and quality problems that caused lower-than-expected yields of defect-free processors when the fabs were completed. Facing ever-increasing pressure to accelerate the design and construction of the fabs even further while maintaining high quality, Art Smith wondered how to redesign Micro's fab engineering and construction work processes and organizations to execute these complex and concurrent projects in a controlled manner.

Art's existing design and construction specialists were organized in a "weak matrix" structure, in which specialists were collocated with their disciplinary colleagues and evaluated by their functional managers to facilitate the sharing of technical best practices. Art considered several options:

- Should he reorganize the team into a strong matrix configuration with dedicated and co-located specialists from all key disciplines reporting to, and evaluated by, a strong project manager? How much time would this save on each project, and how might this change impact the capturing and sharing of technical best practices?
- Should he add additional technical staff and/or substitute higher-skilled engineers or craft workers for those currently on the project team, and if so, for which disciplines or crafts?
- Should he add more management personnel, and if so, where in the team and with what kinds of management skills – schedulers, cost engineers, quality control managers?
- Should he re-sequence tasks to be more or less concurrent? How much time could this save and with what impacts on expected cost and/or quality?
- Should he decentralize decision-making to speed up exception handling? What impact might this have on expected quality?

Art could not find any systematic way to help him make these kinds of decisions. Absent any credible tools for designing his project organization systematically, his default – along with the managers of many other large, complex, and costly projects – had become to treat each multi-billion dollar fab design and construction project as a costly, and potentially career-ending, trial-and-error experiment on the path to discovering a way to optimize the organization and work process for fab delivery.

Design Theory, Methods, and Tools for Physical Systems

The engineers and managers working on the chip design and manufacturing engineering side of Micro operated in a world where the designs of their increasingly complex and densely arrayed microprocessors could be modeled, tested, iterated, and refined in advance, using computational analysis tools to predict the performance of a given case in many different dimensions – e.g., logic validation, spatial layout, induced stray current, heat flow, etc. – with considerable accuracy. This systematic and multidimensional model-based design approach for its products was already well advanced and quite routine. What Micro lacked – and what Art Smith challenged a group of Stanford researchers to develop – was a comparable design theory, methods, and tools that Micro's project managers could use to model and analyze a proposed organization and work process case for a fab's design and construction and predict its cost, schedule and quality performance. This would allow his project managers to iterate through analyses of multiple alternative cases of work processes and organizations conveniently and rapidly, and find a case whose performance would best meet the scope, schedule, and resource objectives for each fab project.

The theory and analysis tools for designing semiconductors – along with bridges, skyscrapers, automobiles and airplanes – rest on well-understood principles of physics and operate on continuous numerical variables describing materials whose properties are relatively uniform and straightforward to measure and calibrate. These physical systems could already be analyzed in the early 1900s by solving sets of linear or differential equations that modeled the components of the physical system and their interaction. Starting in the early 1960s, analysis of these systems was increasingly carried out via numerical computing methods that evolved from the World War II use of computers to calculate ballistic trajectories and crack enemy codes. The approach used to develop the engineering science and technology for analyzing and predicting the behavior of physical systems was to:

1. break a large system into smaller elements whose behavior and interactions could be described;
2. embed well-understood micro-physics theory into the elements;
3. attempt to reflect the interactions between elements through constraints (such as constraints that conserve mass or energy, or that maintain consistency between shared element edges in a finite element structural analysis model); and
4. use the vastly more powerful number-crunching ability of computers (compared to human brains) to simulate the system of elements behaving and interacting under various sets of external loads to predict the element- and system-level behaviors of interest.

The result was that engineers rapidly gained the ability to make increasingly accurate predictions of both micro and macro behavior of many kinds of engineered systems. Some of the earliest pioneers in this computational modeling and simulation of physical systems were civil engineers solving large structural engineering problems. For many kinds of structures, design tools can now predict stresses, strains, and deflections under a variety of loading conditions to finer tolerances than the structure can be built.

Design Theory, Methods, and Tools for Organizations

In stark contrast to the sophistication of engineers in modeling physical systems, theories describing the behavior of organizations are still almost exclusively characterized by nominal and ordinal variables, with poor measurement reproducibility. With very few exceptions, the prevailing theories that could be used to describe or predict the behavior of organizations in the late 1980s were verbal descriptions that incorporated nominal and ordinal variables. Theories expressed verbally using nominal and ordinal variables create a significant degree of linguistic ambiguity, so that results of natural or synthetic experiments cannot always be reliably replicated, and contrasting or competing theories are difficult to reconcile or disprove. Thus, developing a quantitative, model-based theory, methods, and tools for designing organizations and the work processes they execute was a daunting challenge.

A key challenge for more systematic design of enterprise-level organizations is that their goals are often vague, diffuse, and contested (March & Simon, 1958). Consequently, it is difficult to evaluate the outcomes of alternative cases, even if one could predict them. However, within such organizations, a specific project encapsulates a subset of the organization's overall employees or contractors that have been assembled for a relatively well-defined purpose with clear and congruent goals, fixed durations, and clearly defined participants assigned to each of the project tasks. Thus, when faced with the challenge of developing reliable quantitative tools for analyzing the performance of organizations, we believed that the performance of project organizations should be relatively easier to predict and evaluate than the performance of enterprise-level private or public organizations, for which all of these process and outcome variables are much more difficult to identify, measure, predict, and evaluate.

THE BIRTH OF VDT

In the late 1980s, when presented with Art Smith's challenge, our research group had the intuition that it might be feasible to develop computational analysis tools to model and simulate project organizations with reasonable fidelity through the application and integration of two computer science technologies that were just emerging from computer science research laboratories:

1. **Agent-based simulation** (analogous to the finite element modeling approach for physical systems described above) had been pioneered for organizations in the classic garbage-can model of organizational decision-making (Cohen, March, & Olsen, 1972). Agent-based modeling approaches allow modelers to: specify and embed relatively simple behaviors (e.g., processing quantities of information or communicating with other agents) in a set of computational agents; specify and operationalize a few kinds of interactions between agents and tasks; and run the simulation to generate emergent behavior from the micro-behavior and micro-interactions between agents.
2. **Non-numerical, general "symbolic representation and reasoning techniques"** were just emerging from the laboratories of "Artificial Intelligence"(AI) researchers at Stanford, MIT, Carnegie Mellon University, University of Massachusetts, Xerox Palo Alto Research Center (PARC), and elsewhere to represent and reason about nominal and ordinal variables (as well as numerical variables). These new representation and reasoning techniques allow the inheritance of properties from "parent classes" to "child subclasses or instances" of those classes (e.g., from "workers" to "craft workers" to "carpenters" to "Joe the Carpenter"); this allows the creation of prototypical "classes" that encapsulate the attributes and behavior of tasks, workers, milestones, etc. and thus allow the rapid creation of instances of these classes that inherit all of the class properties and behavior and can rapidly be assembled into a realistic model of the work process. These early AI tools like SmallTalk (Goldberg & Robson, 1983), developed at PARC, and Knowledge Engineering Environment (KEE), developed by Intellicorp, a Stanford spinoff, also supported inferential reasoning about the attributes of objects using "If…, then…" production rules and other forms of computational inference.

The Virtual Design Team (VDT) research was thus initiated in 1987 through Stanford's Center for Integrated Facility Engineering with the goal of developing new micro-organization theory and embedding it in software tools. Our intuition was that agent-based simulation using a combination of non-numerical and numerical reasoning techniques could potentially allow us to model and simulate information flow in organizations and the emergent cost, schedule, and resource outcomes of information processing and communication by and between members of project teams. From the beginning the goal was to develop and validate methods and tools to predict the behavior of organizations executing their work processes with both high fidelity and transparency. The fidelity would give managers the confidence to use the methods and tools to analyze, predict, and optimize the performance of their engineering organizations. Transparency would make the tools easy enough to use and understand that managers could begin to use them in the same way that engineers design bridges, semiconductors, or airplanes – by modeling, analyzing, and evaluating multiple virtual prototypes of the work process and organization in a computer, supporting both decision-making and the development of organizational insights. A key early decision was to use professional programmers and develop drag-and-drop graphical user interfaces to support the robustness, ease of use, and transparency of VDT.

The extremely creative and insightful garbage-can model of decision-making developed by Cohen et al. (1972) was an elegant and simple, yet fruitful, agent-based simulation model of university participants engaged in decision-making meetings. The success of this effort persuaded us, along with many other researchers (e.g., Epstein & Axtell, 1996; Masuch & LaPotin, 1989), to explore the use and limitations of agent-based simulation of organizations. The garbage-can model was a relatively abstract, high-level model of organizational decision-making; Masuch and LaPotin (1989) subsequently extended the model and elaborated both tasks and organizational participants to a much finer-grained level of detail that could potentially have been validated against real micro-organizational behaviors and outcomes (although they did not attempt this kind of validation). These two efforts were important points of departure for our research.

GOALS AND PHILOSOPHY OF THE VDT RESEARCH PROGRAM

Note that the goals of the VDT project were different from those of the two models described in the previous section. Previous organizational modeling and simulation researchers had aimed to use simulations to explore, develop, and test new meso- or macro-level descriptive theory, rather than to emulate and ultimately predict micro-reality. An engineering analysis tool emulates the behavior of its physical elements as accurately as possible and predicts the behavior of the elements and the emergent behavior of the larger system to enable prediction, iterative refinement, and consequential interventions in the design of the product or process being modeled. Our goal was to produce an analysis tool that would support the explicit design of particular project organizations containing workers with defined skill sets and experience levels to execute given work processes under specific and tight resource and time constraints. So we needed to quantify the variables in the model and validate the model's micro-behaviors and predictions extensively for it to become useful for our intended purpose.

By predicting the performance of alternative configurations of an engineered system, model-based simulation can provide engineers or managers with the ability to conduct multiple "virtual trial and error experiments" in which they test – and often "break" – virtual rather than physical prototypes of candidate solutions. Thus, if the modeling methods and tools are easy and transparent enough for managers to develop and explore multiple configurations in a reasonable amount of time, the managers can develop tacit knowledge and expertise about the performance contours of different configurations of a proposed solution by experiencing how the different configurations break in different ways. Accordingly, we decided to call our engineering project modeling and simulation system the "Virtual Design Team" (VDT), by which we meant a computer simulation model of a real design team.[2]

Direct Work and Three Kinds of Hidden Work

VDT was based on the notion, articulated by Herbert Simon (1947), refined by Jay Galbraith (1974), and extended and quantified by our research team, that the first-order determinant of an organization's success is its ability to process all of the information associated with *direct work* as individuals or groups complete their assigned tasks; and *exceptions* arising from missing or incomplete information needed by a worker to complete an assigned task. Each exception requires the worker to seek advice from a more knowledgeable person, generally a supervisor somewhere up the hierarchy. Galbraith had proposed this idea as early as the 1960s, but his formulation of the problem was descriptive and qualitative and thus could not be used to make specific predictions about when and where the quantity of information to be processed in a specific work process would overwhelm one or more participants in the organization assigned to execute that work process. VDT quantified, extended, and validated Galbraith's information-processing view of organizations conducting work and generating, escalating, and resolving exceptions to encompass a broad range of project-oriented work processes and organizations. In refining and elaborating Galbraith's notion of exceptions, we distinguished between:

- ***Functional exceptions*** arising from incomplete technical knowledge, which a worker might escalate to a more expert functional supervisor in his or her discipline who would be required to do "supervisory work" to resolve the exception
- ***Project exceptions*** arising from incomplete information at the interfaces between interdependent tasks performed by peers in other disciplines, which a worker would need to resolve by doing "coordination work" with the interdependent party – what Thompson (1967) referred to as "mutual adjustment of reciprocal interdependency"
- ***Institutional exceptions***, arising in cross-cultural global project teams from the need to resolve differences in goals, values, and cultural norms between project team members from different national institutional backgrounds (Scott, 2008). Managers attempting to resolve this kind of exception would need to perform "institutional work." We set institutional exceptions aside for subsequent research and focused initially on modeling functional and project exceptions.

2 The phrase "virtual team" subsequently began to take on a different colloquial meaning in the organizational literature – a geographically distributed team and/or one comprised of members from multiple separate organizational entities.

The intuition behind the 20-year VDT research program was that direct work, supervisory work, coordination work, and institutional work could all be viewed as quantities of information to be processed by the workers and managers in an organization. If one could represent and quantify the information-processing demand generated by a given work process, and the information-processing capacity of the workers and managers in an organization configured in a particular way, a simulation model of the flow of information to perform direct work and generate and handle exceptions through a project team would provide a first-order estimate of whether or not a given configuration of the project organization possessed the appropriate information-processing capacity in the correct places within the project organization to:

- process the information required to execute the direct tasks;
- provide adequate, high-level technical information-processing capacity in the right places to resolve technical exceptions; and
- have sufficient slack information-processing capacity to allow interdependent workers to coordinate cross-disciplinary reciprocal interdependencies that might arise in the execution of the project.

In this respect, VDT is simply a micro-level, more detailed and quantified form of the qualitative, rule-based macro-information processing contingency theory framework used to diagnose organizational misfits in Burton and Obel's (2004) book *Strategic Organizational Diagnosis and Design* and its accompanying *Organizational Consultant* software tool.

Organizational Physics, Chemistry, and Biology

We viewed this analysis of the project organization's information-processing capacity vs. information-processing demand as a first-order "information flow physics" approximation of the organization's ability to execute the project. In this respect, VDT is similar to Isaac Newton's Second Law of Motion, which predicts the motion of an object subject to one or more force vectors – but without considering effects like friction or relativity – accurately enough for many practical purposes. If the physics of a bridge are inadequate, it collapses the first time the wind blows too hard, like the first Tacoma Narrows Bridge. Similarly, if the information-flow physics of a project organization are wrong, the organization encounters cost overruns, schedule overruns, and quality risks in a way that Galbraith predicted qualitatively from his observations of aerospace projects in the 1960s. VDT assumes uniform and high levels of motivation by all project actors and ignores the potential for goal conflict. A more refined analysis of the goals and motivation of actors – which we excluded from our first-order physics model – can be viewed as "organizational chemistry." If the organizational chemistry is wrong, the organization eventually fails through slow processes analogous to "corrosion" of physical systems. Finally, if the "organizational biology" is wrong, the organization cannot grow new knowledge to enhance its performance over time or reproduce itself.

As we discuss later in this article, subsequent versions of our VDT model began to incorporate some aspects of organizational chemistry and organizational biology. This paper will focus primarily on the information flow physics of our first VDT prototype, "VDT-1."

VDT MODELING AND SIMULATION APPROACH

We directed our initial focus toward project organizations engaged in semi-custom engineering work under tight time constraints, such as those encountered by Micro in our example above. For such organizations, we could assume a relatively high level of congruency of goals, culture, and values, so that institutional work is negligible and can be ignored. However, performing highly interdependent work under tight time constraints creates a significant amount of coordination work as interdependent tasks increasingly overlap one another in time. Primary emphasis was on modeling the sources of interdependence in project workflow and the way in which exception handling and coordination took place within organizations assigned to do such project work.

VDT incorporated the kind of quantitative reasoning about decision-making demand and capacity used in the garbage-can model (Cohen et al., 1972) as well as the kind of non-numerical reasoning about task assignments, skill sets of participants, etc. used in Masuch and LaPotin's (1989) model. VDT uses symbolic reasoning about nominal and ordinal variables (e.g., the degree of fit between the worker's skill set and skill level vs. the technical complexity and uncertainty of the task to which the worker is assigned) to set parameters for numerical variables (e.g., task processing speeds and expected error rates) in a quantitative, stochastic, discrete event simulation. In the remainder of this section we provide an overview of the representation and reasoning in VDT.

Modeling a Project in VDT

A VDT user assembles a work process and organization configuration (called a "case" in VDT) using a graphical "model canvas" to provide maximum transparency of the modeled case for the manager and model developer.

- **Project organization participants** are rapidly created by dragging and dropping team members from a graphical palette onto the model canvas as instances of classes defining the behavior of three kinds of employee roles (project managers, sub-team leaders, or sub-teams).
- Similarly, specific **tasks**, milestones and meetings are created as instances of classes (e.g., milestones, tasks, and meetings) by dragging and dropping the appropriate objects from the palette onto the model canvas.
- Several kinds of **relationships** between actors and other actors (i.e., supervisory relationships), between pairs of tasks (e.g., sequential interdependence, information exchange requirements), and between actors and assigned tasks (e.g., primary or secondary task assignments, meeting participation) are created by dragging and dropping relationship objects from the palette onto the model canvas and connecting them between the appropriate actors or tasks.
- **Contextual variables** such as overall project complexity and uncertainty, the strength of the functional vs. project dimensions of the matrix, the prior experience of team members working with one another, etc. are entered into a property table prior to simulation.
- **Agent micro-behaviors** for different types of work – e.g., hardware engineering vs. software engineering – are defined using a set of small matrices stored in a "behavior file." The rows and columns in these behavior matrices are typically nominal or ordinal variables that describe actor, task, or context properties – e.g., an actor's *Application Experience* (the level of experience the actor has working on this type of task, with values of low, medium, or high) and the actor's *Skill Level* in the profession involved (say *Structural Engineering*, rated as low, medium, or high). The entries in each cell of this 3x3 matrix are numerical values used in the discrete event simulation, e.g., a number that is the ratio of the actor's information-processing speed relative to a nominal actor who has medium application experience and medium skill of the type required to perform this task. In our research we developed and validated two predefined behavior files: the default behavior file developed from construction, aerospace, and other kinds of hardware engineering; and a second optional behavior file with significant differences that more accurately describes agent micro-behavior for software engineering. These matrices are contained in a text file and can easily be edited and modified to model different kinds of agents engaged in other kinds of work processes. The ability to edit the behavior files easily has been exploited by many of the researchers whose experiments are described in the section on "Using VDT to Develop Meso- and Macro-Organization Theory."

The VDT model canvas for the project manager's initial "Baseline Case" of the work process and organization to complete the design of a biotech manufacturing plant is shown in Figure 1.

Simulating Project Organizations in VDT

The Virtual Design Team simulation system is an agent-based, computational, discrete event simulation model of information flow in project organizations. As VDT actors attempt to complete their direct work, task attributes such as complexity and uncertainty and actor attributes such as skill level and experience are evaluated and compared. VDT reasons qualitatively about non-numerical attributes such as individual team members' skills and experience, task attributes like work volume, complexity, and uncertainty, and ordinal organizational variables such as the level of centralization and formalization (high, medium, or low) to set numerical values like actor information-processing speeds, and exception rates for functional and project exceptions used in the quantitative discrete event simulation. VDT simulates each of the team members processing its assigned tasks, once the tasks' predecessors have been completed, and generates functional and project exceptions stochastically using Monte Carlo sampling methods.

Actors are more likely to generate exceptions when confronted with a task for which they do not possess the requisite levels of skills or experience. Depending on the advice of the manager to whom an exception was delegated, the actor may need to rework the task that generated the exception partially or completely. Actors may be required to attend to communications from other actors and may need to attend scheduled meetings, all of which consume the actor's information-processing capacity. Moreover, failure of an actor to attend to a communication within a specified length of time (after which the communication is moot) or to attend an assigned meeting increases the probability of exceptions occurring downstream. These kinds of communication failures thus produce second-order effects such as increased downstream coordination and rework costs.

A detailed explanation of the objects, attributes, relationships, and behavior in VDT is beyond the scope of this paper. Interested readers are referred to Jin and Levitt (1996).

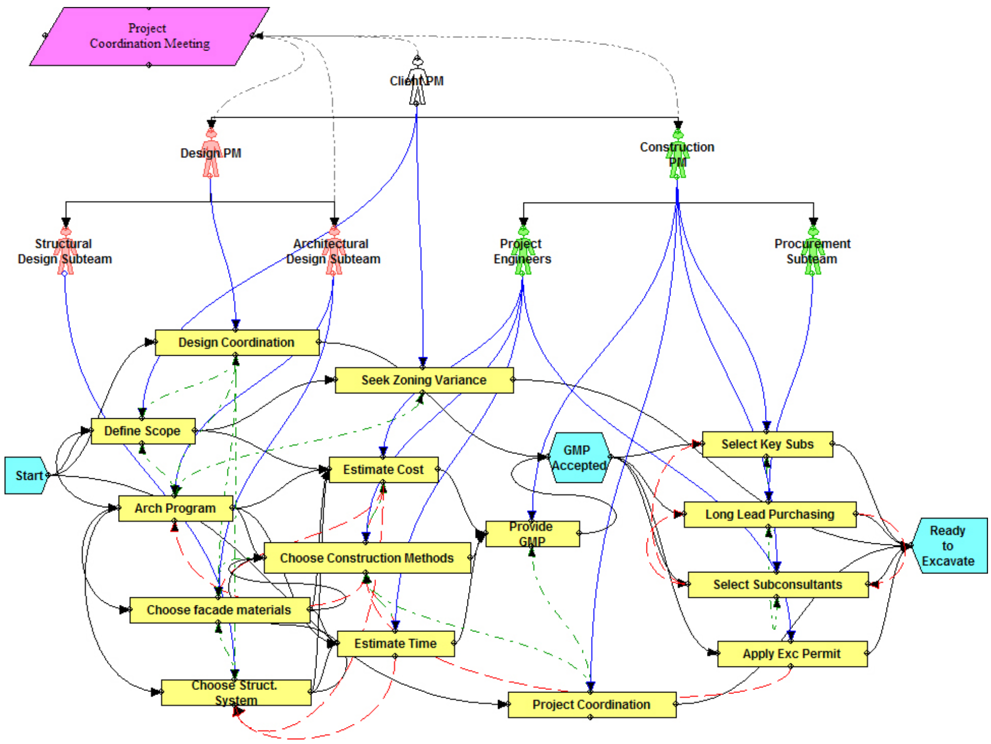

Fig. 1. VDT/SimVision Graphical Model Canvas

VDT thus builds on and quantifies Galbraith's (1974) information-processing view of project teams and views both the direct work and resulting coordination work on a project as quanta of information to be processed by assigned actors who have only "boundedly rational" (March & Simon, 1958) information-processing capacity. It simulates the project team executing tasks and coordinating to resolve exceptions and interdependencies. The simulation of a project organization executing its tasks generates a range of outputs that predict the emergent performance of the organization at both the individual actor/task level and the overall project level: duration, production costs, coordination costs (communication, rework, waiting), and several measures of process quality.

Iteratively Refining a Project's Organization and Work Process Using VDT

The approach used by a manager like Art Smith to design an organization using VDT starts by having the manager generate a plausible first cut at the organization and work process for his or her project based on his or her prior project experience and/or judgment. The manager can then simulate this first cut "Baseline Case" to see how well its predicted schedule, cost, and quality risk meet project goals. Figure 2 shows a Gantt chart to visualize the predicted schedule performance of the baseline organizational case for the biotech design project shown in Figure 1. The Gantt chart shows this biotech project will achieve its completion milestone of "Ready to Excavate" (black diamond on the last line of the Gantt chart) in early March of 2007, long after its planned early December completion date (green diamond on the final line).

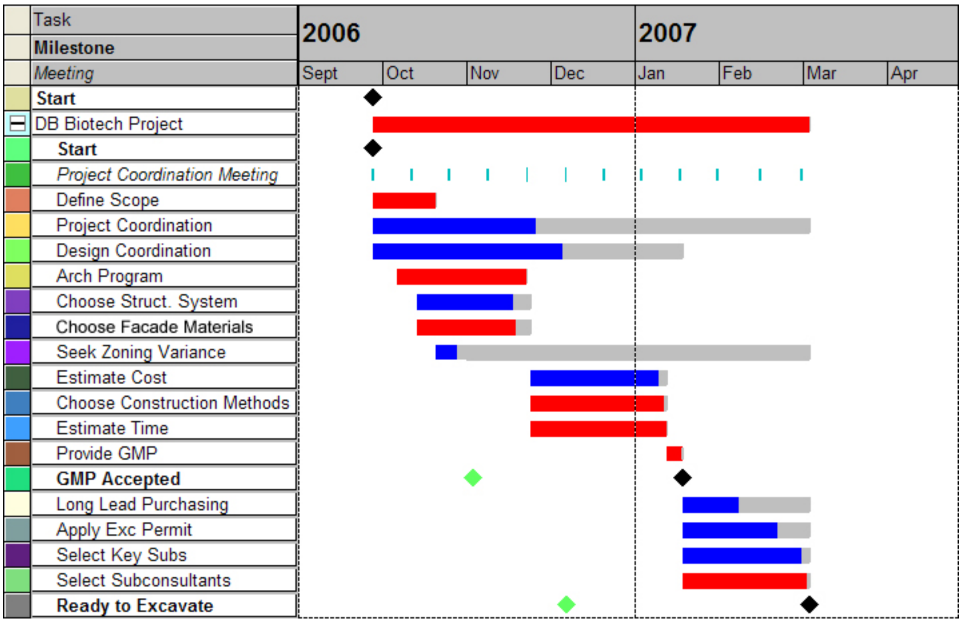

Fig. 2. VDT/SimVision Simulation Schedule Output

The VDT model canvas[3] shown in Figure 1 was used to create and visualize the work process and organization model for a project to accelerate the design of a biotech manufacturing plant for a recently approved cancer therapy drug. Tasks, milestones, and organizational participants are dragged and dropped from the model palette on the left onto the canvas and named. They can then be connected into relationships such as: task-activity successor links, shown as black arrows; the supervisor-subordinate hierarchical relationships shown in the project organization chart; or the blue task assignment links between participants and their assigned tasks by dragging and dropping the appropriate connector onto the model canvas and connecting the ends to the attachment points on the desired objects. The purple object at the top left is a weekly two-hour coordination meeting, attended by the project manager and sub-team leaders connected to it with dashed arrows. Numerical project-level parameters for technical and cross-functional error probabilities, information exchange frequency and noise, and low, medium, or high ordinal values for organizational parameters such as matrix strength, team experience, centralization, and formalization are entered directly into the property table at the top left. Clicking on any object displays its properties (e.g., team members' skills and skill levels, tasks' total work volume, etc.) in the property pane, where they can be input and changed.

3 VDT was commercialized in 1996 as SimVision™. The VDT modeling canvas was a slightly more primitive, but essentially similar, version of the SimVision modeling canvas shown in Fig 1. (SimVision is licensed by ePM of Austin, Texas http://epm.cc for academic use or professional application).

If this were his project, Art would want to understand why the project was predicted to be so late. The bars shown in red on the Gantt chart indicate critical path[4] tasks whose duration determines the final completion. Blue bars with gray "float" shown after them are non-critical tasks whose duration will not impact project completion. It would be helpful if Art could determine which organizational participants were predicted to be backlogged with information overload in the baseline case. Figure 3 shows the VDT prediction of the Information-Processing backlogs in Full-Time Equivalent (FTE) person-days for all of the positions in the project organization.

Art could then make up a second project case to explore the implications of an intervention such as: increasing the capacity of one or more of the most heavily backlogged sub-teams *(Architectural Design Team and Construction PM)* responsible for tasks that lie on the critical path; increasing the skill level of the workers already assigned to those tasks (by substitution of more experienced team members or training of existing team members); changing the sequence relationships between tasks on the critical path so that they are performed concurrently rather than sequentially; etc. He could then simulate this second case to evaluate its performance in terms of project objectives, and compare its performance to the baseline case to see whether this intervention to the baseline case predicted a better or worse trade-off among his project objectives. Figure 4 compares the schedule for an intervention that adds 0.5 FTE to the Architectural Design Team and 1.0 FTE to the Construction PM to the Baseline Case.

This figure shows the VDT schedule prediction for the Baseline Case of the biotech plant example shown in Figure 1. The client wanted the project to be ready for construction by the first week in December – the green "Planned Milestone Date" diamond on the final Ready to Excavate row of the Gantt Chart – in order to get the foundation built before the rains begin. VDT predicts that the Baseline Case will be completed in mid-March, about three months late, shown by the black "Predicted Milestone Date" diamond at the lower right. This is clearly an unsatisfactory case, so the manager will need to model and simulate possible interventions in the project scope, work process, and/or organization to find a case that will allow him or her to complete this project on time.

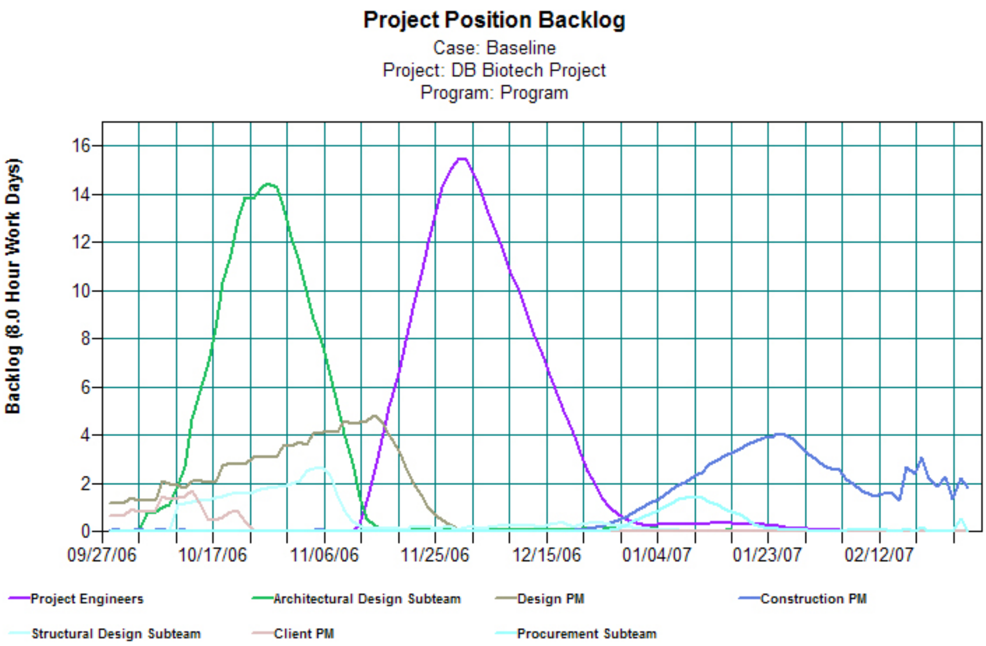

Fig. 3. Predicted Information-Processing Backlogs

This chart shows VDT's predictions of the expected full-time equivalent (FTE) person-days

4 The "critical path" is the path through the longest chain of sequentially dependent tasks in the project. The durations of activities that lie along the critical path determine the project duration, since any change in the duration of one of these tasks will impact the final completion date of the project.

of backlog for all of the positions shown on the organization chart in Figure 1. Note that the Architectural Design Team is predicted to be backlogged about 14 FTE-days early in the project and the Construction PM is predicted to be even more backlogged in the latter part of the project. When backlogs get beyond about two FTE days, managers focus on recovering from their own backlog of direct work and may fail to respond to coordination requests before they time out and miss scheduled meetings, causing quality risks to rise. Adding extra capacity or raising the skill levels of the persons assigned to one or both of these two positions will likely improve the schedule and may also have implications for the project's process-quality risks.

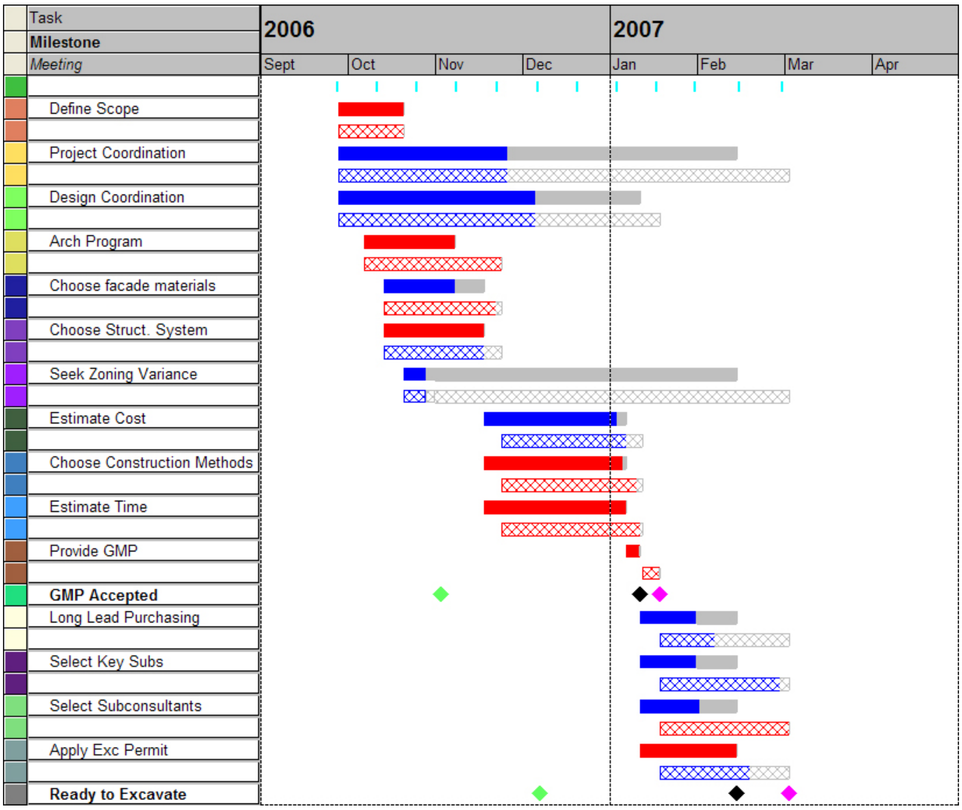

Fig. 4. Exploring the Impacts of an Intervention on Project Schedule

This Gantt chart shows the effect of adding 0.5 full-time equivalents (FTEs) staffing to the *Architectural Design Team* and 1 FTE to the *Construction PM*. The task durations and start and end times for the modified case are shown as solid bars and can be compared to the original Baseline Case shown as hatched bars; the milestone dates for the new case are shown as black diamonds, and those for the Baseline Case are shown as purple diamonds; the client's planned milestone dates are shown as green diamonds. A glance at the bottom line – the Ready to Excavate completion milestone – shows that this intervention will shorten the project by about three weeks from the Baseline Case, but will still complete much later than the planned completion date (the green diamond on that row of the Gantt chart). Scanning the bars to see where the time savings were achieved and where the critical path now lies reveals that the biggest impacts of this intervention case were to shorten the duration of the two critical path tasks, *Arch Program* and *Choose Façade Materials*, performed by the *Architectural Design Team*. Note that *Choose Façade Materials* is now predicted to be non-critical. Similarly the durations of the tasks, *Select Key Subs* and *Select Subconsultants*, performed by the *Construction PM*, have been shortened. *Select Subconsultants* was previously on the critical path, but both tasks are now non-critical.

Thus far, we have only considered schedule goals; a more thorough analysis must also assess whether desired cost and quality metrics have been achieved. These outputs are shown schematically at the right of Figure 5. Unacceptable performance in terms of cost or quality risks can be addressed by different kinds of managerial interventions. For example, unacceptably high levels of functional quality risk can usually be addressed by increasing the level of centralization of decision-making to *High* (i.e., most exceptions will now be reviewed by project managers instead of sub-team leaders). However, this can introduce delays if a backlogged project manager takes longer to attend to, and resolve, exceptions. Organizational contingency theory (Burton & Obel, 2004) asserts that this trade-off depends on several contextual variables, such as the span of control of the project organization (how many sub-team leaders report to the manager, and how many workers report to each sub-team leader). The higher the span of control at each level, the larger the number of workers reporting to that manager, and hence, the greater the expected frequency of exceptions landing in the managers' in-basket. If the project organization has a high level of centralization – i.e., most exceptions must be dealt with by the project manager – then a large span of control, coupled with a relatively poor match between the workers' skills and the complexity of the tasks they are working on, will result in a high likelihood that the project manager will get backlogged and become very slow to handle exceptions.

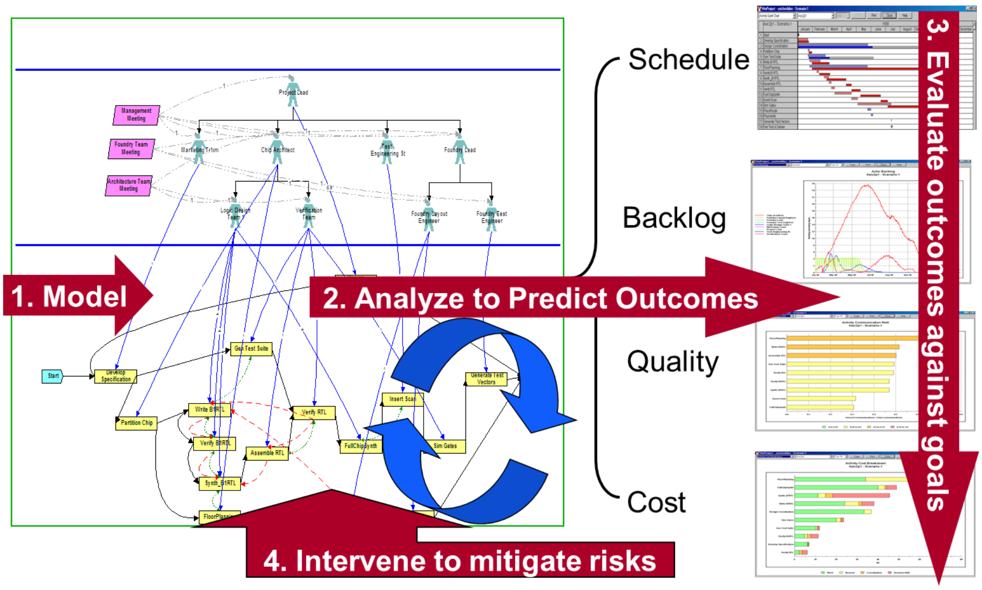

Fig. 5. A Process Model for Simulating and Evaluating Project Outcomes

High backlogs do not only affect project schedule. When managers become backlogged and fail to handle exceptions within a reasonable timeframe, subordinates begin to "delegate by default" – i.e., they use their best judgment to decide what to do about an exception. When this occurs, the level of centralization of decision-making in the organization has effectively been lowered by default rather than by design. VDT models these "delegation by default" instances as increasing the "functional quality risk" for the tasks whose exceptions have been delegated by default to low levels of decision-making.

Similarly, cross-disciplinary coordination can break down if workers who are asked to respond to coordination messages fail to respond within a reasonable period, resulting in increased "communication risk" for the task whose coordination was not completed. Unacceptably high communication risk can be addressed by increasing the project organization's matrix strength. This is achieved in practice by co-locating team members of different functions in a project cluster and having the project manager evaluate them in terms of project objectives rather than having a functional manager evaluate them based on each discipline's technical criteria. Note that increasing the organization's matrix strength will decrease communication quality risk, but it can increase technical quality risk because functional workers are no longer co-located with their functional peers.

These are precisely the kind of difficult and opaque organizational trade-offs that can be explicitly and transparently explored by a manager using VDT/SimVision. A quantitative simulation tool like VDT/SimVision provides **quantitative resolution** of the **qualitative**

indeterminacy that is otherwise inherent in these trade-offs. Proceeding iteratively in this way, the manager can explore the implications and trade-offs among schedule, cost, and quality outcomes resulting from dozens or even hundreds of alternative cases of the organization and work process in order to find one or more alternative cases that come closest to meeting project goals. If the project goals cannot be achieved through changes in the work process or organizational structure – which is often the case for projects with very aggressive schedule goals – the manager can explore reducing the scope of the technical deliverables for the project. In many cases, it may be more advantageous to the client to scale down the project's scope in ways that do not detract from its primary function in order to have at least a scoped-down version of the product ready by a fixed date such as a tradeshow or a regulatory deadline. This will shorten task durations and possibly eliminate some tasks, positions, and/or staff members from the project team. In the biotech design case illustrated above, the client ultimately found that the desired early December completion date could not be met with any feasible configuration of the work process or organization, and therefore decided to use a prefabricated metal building for the biotech facility instead of having the architect design a custom building for the plant. This greatly reduced the scope of the architectural design tasks and resulted in a predicted early December completion date, which the team was able to meet.

The process of modeling, simulating, and evaluating predicted outcomes against project goals, and iteratively refining and testing alternatives in an attempt to better meet project goals, is summarized in Figure 5. By iteratively modeling, analyzing, and evaluating alternatives, and exploring the impact of successive interventions, a manager can rapidly explore dozens or hundreds of cases of the work process and organization, and home in on one or more cases that provide the best trade-off among scope, schedule, cost, and quality project objectives.

VALIDATION OF VDT

In their paper on validation of computational organizational models, Burton and Obel (1995) cite Cohen and Cyert (1965), who asserted that "...even though the assumptions of a model may not literally be an exact and complete representation of reality, if they are realistic enough for the purposes of our analysis, we may be able to draw conclusions which can be shown to apply to the world." Thus, some models must be rather realistic; some need not be. As explained above, the primary goal of the VDT research was to develop a computer simulation model that could emulate the behavior and outcomes of real-world project teams executing complex work processes accurately enough to guide managerial interventions. Thus, it was important to us that we carefully validate and calibrate the non-numerical and numerical parameters of the model's inputs and outputs so that we could eventually credibly claim that VDT provides accurate first-order predictions for real-world projects.

By operationalizing and extending Galbraith's information-processing abstraction in the VDT computational model, and focusing on semi-routine project organizations – an "easy corner" of the space of all organizations – we developed several versions of VDT and validated the representation, reasoning, and usefulness of our computational "emulation" models using the rigorous validation trajectory shown in Figure 6 (Kunz, Christiansen, Cohen, Jin, & Levitt, 1998; Levitt, Cohen, Kunz, Nass, Christiansen, & Jin, 1994; Levitt, Thomsen, Christiansen, Kunz, Jin, & Nass, 1999; Thomsen, Levitt, Kunz, Nass, & Fridsma, 1999). The large background arrow charts the validation trajectory from the lower left to the upper right of this diagram, showing how we successively validated the Reasoning, Representation, and, finally, Usefulness of VDT.

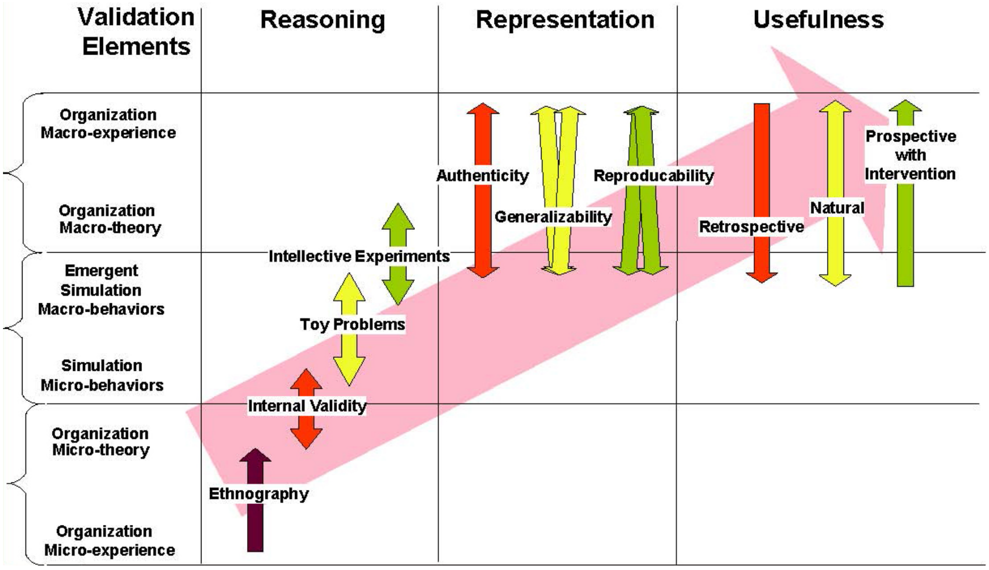

Fig. 6. Validation Trajectory for VDT Project Organization Simulation Model
Source: Thomsen et al., (1999).

Validation of Reasoning

Phase 1 of the validation focused on the model's "reasoning" – the parameters and algorithms that simulate information-processing and exception handling by agents in the model. This phase required, first, that the micro-behavior of workers and managers in the model be based on solid ethnographic research by our research team or others. Thus, we began our research in 1988 by using ethnographic methods involving shadowing of project team members and their managers for weeks at a time to gather quantitative data on low-level actor and task behaviors, such as the length of time it typically takes managers at different levels to resolve exceptions, the rules project team members use for deciding the order in which to attend to items in their in-baskets, the effect on project error rates of missing meetings, and so on. This ethnographic research was reported in Cohen (1992) and Christiansen (1993). Next, we needed to validate the accuracy of the model's predictions. To do this, we embedded these validated chunks of agent micro-behavior in the simulation agents and designed a set of "toy" problems – small idealized cases involving a handful of tasks and positions for which we could determine the correct outcomes by hand calculation – to validate that we had correctly embedded these behaviors. The third step in evaluating the reasoning (Christiansen, 1993) was to design "intellective" experiments (Burton & Obel, 1995) in which we attempted to replicate the predictions of information-processing organization theory developed by others, drawing on the encyclopedic compilation of organizational contingency theory in Burton and Obel (2004).

"Docking" two or more computational models of organizations against the same set of data to compare their outcomes has been proposed as a particularly insightful form of validation of the respective models' reasoning. Several researchers have used VDT/SimVision in docking experiments with Burton and Obel's (2004) OrgCon, including the following: Carroll, Gormley, Bilardo, Burton, and Woodman (2006) docked SimVision against OrgCon to study project work processes and organizations at NASA, yielding valuable insights for the NASA managers; and Cardinal, Turner, Fern, and Burton (2011) carried out an ambitious experiment involving a three-way triangulation of SimVision against both OrgCon and data from a set of case studies of new product development, and were able to develop new contingency theory propositions for the design of product development organizations. Similarly, Carroll and Burton (2012) carried out a three-way docking using SimVision to optimize project organization design; OrgCon to diagnose the goodness of fit of the elements of NASA's enterprise's organization and context; and the *Design Structure Matrix* tool (Steward, 1981) to analyze task interdependence and reorder tasks to minimize design cycles. Each of these experiments demonstrated the feasibility of using multiple organizational analysis tools side by side to design project organizations, and they highlighted the complementarity of the tools involved for shedding light on different aspects of the design of the project organizations and their work processes.

Validation of Representation

The second phase of the validation assessed VDT's semantics and syntax in terms of their "representational validity." This consisted of validating its "authenticity" – i.e., whether the terminology in VDT was easily and consistently understood by practitioners – the "generalizability" of the VDT modeling concepts across different kinds of projects, and the "reproducibility" of models – i.e., whether different modelers would produce similar VDT models of the same project. Cohen (1992), Christiansen (1993), and Thomsen, Kwon, Kunz, and Levitt (1997) all contributed to this phase of the validation by working with managers of real projects and observing when names of objects, relationships, or other model inputs and outputs did not match the manager's colloquial understandings of those terms (e.g., we changed the nomenclature of "Role" to "Position"; "Actor" to "Person"; "Activity" to "Task"; "Exception" to "Error"; etc., as a result of our validation of the model's authenticity. We modeled several different kinds of engineering projects, including oil refineries, electric power substations, biotech manufacturing plants, semiconductor fabs, software development efforts, satellite launch vehicles, satellites, and microprocessors in different phases of the validation. In addition to the research students who formally validated the representation, reasoning, and authenticity of models, about 50 MS-level graduate students per year over a period of about eight years used our evolving VDT modeling and simulation methods and tools in project organization design classes in which they modeled more than 100 other projects in a variety of different domains and provided valuable feedback to the research team on representational issues.

Validation of Usefulness

The final phase of the validation focused on the model's "usefulness" – the extent to which project management practitioners would eventually come to have enough confidence in VDT's predictions to begin using the model to support organization design proactively on their projects. This phase involved modeling and attempting to emulate the outcomes of real-world projects – first retrospectively, then in real-time natural experiments. Cohen (1992) retrospectively modeled the repairs to a series of electrical substations damaged by the 1989 Loma Prieta earthquake that had to be urgently repaired, and adjusted numerous parameters of the model to replicate this past experience. Christiansen (1993) carried out additional retrospective validation of the model's predictions, in which he replicated the design of the Statfjord subsea oil modules that had been designed and installed under extreme time pressure in Norway's North Sea oil fields and calibrated the model parameters associated with quality risks.

Thomsen (1998) conducted the first real-time validation of VDT on Lockheed's attempt to build its first commercial satellite launch vehicle. Lockheed had been building roughly comparable launch vehicles for military missiles for more than two decades, so they viewed this project as semi-routine at this point. However, to meet the needs of very demanding clients, they were attempting to develop a commercial satellite launch vehicle in just one year – one fifth of the time that it had historically taken the company to develop comparable launch vehicles for Navy missiles. The VDT research team was asked by the National Science Foundation, which had provided the bulk of the funding for the VDT research, to study the Lockheed Launch Vehicle One (LLV1) project in real-time and predict its outcome. The project commenced in early 1995 and was scheduled to be completed and launched by the end of that year.

By March of 1995, a team consisting of Jan Thomsen, John Kunz, and Yul Kwon developed a VDT model of the organization and work process for this project and ran the simulation. The simulation predicted that LLV1 would not be completed until mid-April of 1996. Moreover, the VDT model of LLV1 predicted extremely high quality risk for the cable harnesses, a component which Lockheed had decided to outsource to an East Coast company in order to develop its capability for "agile manufacturing" and to save a modest amount of cost.

The launch vehicle was completed and launched about four months late (within a few days of the date VDT had predicted a year earlier). The launch vehicle almost immediately "departed controlled flight" and had to be detonated by the Air Force safety officer. Analysis of telemetry data from the failed launch vehicle indicated that the most likely cause of failure had been a cable from one of the cable harnesses that had been misrouted and got too close to a hot area of the launch vehicle, which melted its insulation and caused a short-circuit – a literal and figurative quality meltdown! As a senior Lockheed manager stated, "The launch vehicle was insured; the satellite was insured; everything was insured except Lockheed's reputation" (Thomsen et al., 1997).

At the time that the Stanford VDT team made its prediction of the completion date and quality risks for LLV in March of 1995, neither they nor the Lockheed managers involved had sufficient confidence in the VDT predictions to intervene proactively in the organization or work process. This extraordinarily accurate natural experiment to predict the outcomes of a real-time project organization was thus a breakthrough moment in the validation of VDT. After this validation exercise, the VDT research team was invited to work with the manager of a subsequent Lockheed satellite project in a different division of Lockheed. This manager helped to build the model and relied on the model's predictions to make a series of prospective managerial interventions that helped keep that project on schedule and within quality bounds (Kunz et al., 1998).

Other researchers subsequently began to use VDT in an "action research" mode for prospective design of project organizations in real-world situations. Carroll et al. (2006) utilized SimVision along with other approaches at NASA to predict project performance, diagnose project risks, and support organizational redesign. This project had a happier – if much less dramatic – ending. Several lessons were learned from this experiment:

- First, similarly to Lockheed's managers, the intuitions of the professional engineers at NASA about the outcomes of alternative project organizations designs was not as good as they believed; their solution was shown to be infeasible using the tools of organizational analysis.
- Second, NASA avoided some headaches and retrofitting that it would have incurred without the tools and their application. That is, NASA avoided an opportunity loss.

Tools can make a difference in the analysis of organizational configurations that have already been designed using managers' intuitions and judgment, or have been copied exactly from previous projects. They can also be used in the upfront design of a baseline organization. The NASA project was a very complicated multi-organizational, multi-location project design where the tools helped managers avoid adverse outcomes.

As Michael Schrage (2000) describes in his book, *Serious Play*, creating a shared language and a visual "blackboard" with which project team members can explore and discuss alternative configurations is valuable in facilitating brainstorming and analysis, even absent any predictive power of the language and visualizations being used. However, when tools like spreadsheets or organizational simulations are able to make plausible predictions about financial outcomes or project organizational outcomes, respectively, the team's decision-making process is literally transformed to a new and much more productive level of brainstorming and decision-making, which Schrage calls "serious play."

Starting in about 1996, after the VDT software had been commercialized as SimVision, consultants at Vité Corporation (the company which initially developed the SimVision prototype under license from Stanford University) and subsequently ePM, LLC, which acquired the rights to the SimVision software and began using the software in its project organization design consulting practice in about 2000, have modeled hundreds of real-world projects with very demanding clients and have demonstrated the usefulness of this model in practice over more than a decade.

By rigorously validating every aspect of VDT in these three ways through all of these validation steps, we were able to generate sufficient confidence in the predictions of our theory and tools that managers in several companies and governmental agencies began using the software to design or redesign their project work processes and organizations prospectively, based on the predictions of this organization modeling and simulation design approach. Our VDT theory and analysis tools for project organizations had thus begun to enable true "organizational engineering" of project organizations that could be assumed to have relatively congruent goals, and were executing relatively routine – albeit complex and fast-paced – engineering-design and product-development work processes.

USING VDT TO DEVELOP MESO- AND MACRO-ORGANIZATION THEORY

Once VDT had been thoroughly validated, researchers at Stanford and elsewhere began to use the simulation tool as a new kind of virtual synthetic organizational experiment to develop, validate, and extend organization theory.

Toward an Organizational Reynolds Number

The first effort of this type was a project that involved several undergraduate students over a number of years attempting to develop an organizational analogy to the dimensionless Reynolds Number[5] that characterizes fluid flow as laminar vs. turbulent in fluid mechanics. Our intuition was that a similar dimensionless number might be found for demarcating the boundary between laminar vs. turbulent flow of information through project organizations based on variables like the span of control of the organization, the degree of complexity of its tasks, and the level of centralization. This kind of Organizational Reynolds Number would then predict the point at which information flow in an organization becomes severely enough bottlenecked that exceptions would generate rework faster than it can be effectively completed (damped out, so that rework generates new exceptions and yet more rework). Exceeding such an "Organizational Reynolds Number" would cause hidden work and project duration both to increase dramatically. Michael Fyall, William Hewlett III, Per Bjornsson, and Tarmigan Casebolt all worked on this research at different times and began to home in on a set of variables that begin to predict when increasing any of these variables would make the information flow become "turbulent" – i.e., it would cause hidden work and project duration to increase exponentially rather than linearly (Levitt, Fyall, Bjornsson, Hewlett, & Casebolt, 2002). This is a truly exciting research challenge that begs for additional research.

Using VDT to Study Knowledge Flows

VDT was subsequently used to develop theory about knowledge flows through organizations by Nissen and Levitt (2004). Nissen and colleagues worked on several different aspects of knowledge flow including the impacts of discontinuous membership in project teams due to turnover or fragmentation across project phases (Ibrahim & Nissen, 2007). Following up on Nissen's work, Levine and Prietula (2011) studied circumstances under which knowledge transfer within organizations would be helpful vs. harmful to the organization.

Exploring Virtual Organizations and the Edge of Chaos

Rich Burton and his students and colleagues have used VDT extensively over the last decade to explore a number of organization theory questions. Timothy Carroll and Rich Burton conducted experiments to explore the "Edge of Chaos" – similar in some ways to the Organizational Reynolds Number work described above (Carroll & Burton, 2000). Zse-Zse Wong and Rich Burton (2000) used VDT simulations of different aspects of virtual organizations – project organizations whose participants were separated by geography and other kinds of distance – to develop propositions about their performance in different contexts. Jensen, Håkonsson, Burton, and Obel (2010) have further elaborated this

5 The Reynolds Number is a dimensionless number that demarcates the boundary between laminar and turbulent flow of fluids. For fluid flowing through a pipe, when the Reynolds Number is below 2300, eddies that are created in the fluid get damped out by its viscosity. For Reynolds Numbers above 4000, eddies begin to generate secondary eddies faster than they can be damped out and the flow becomes turbulent. When the flow becomes turbulent, the pressure loss from fluid flowing through the pipe begins to increase with the square of the fluid's velocity rather than linearly with its velocity. In between these two values, the flow is "transitional" and can be either laminar or turbulent.

research. Kim and Burton (2002) used VDT simulations to study how task uncertainty and decentralization affect project team performance. And Burton and Obel (2011) show how VDT simulations can be triangulated against other simulations and empirical data to extend and refine organization theory. The citations over time for the experiments described above show that publications describing research using agent-based modeling tools like VDT to develop and extend organizational theory have moved from specialized journals focused on computational simulation to mainstream organization theory journals in the last few years.

EXTENSIONS TO THE ORIGINAL VDT MODEL

Since the mid-1990s, Stanford researchers have extended the representation and reasoning in VDT step-by-step, to address the modeling requirements of less routine work performed by increasingly flexible and dynamic organizations – non-routine product development, service and maintenance work (including healthcare delivery), and highly non-routine work performed in communities of practice – but still assuming negligible institutional work. Starting in 2002, we extended VDT to model multicultural project teams engaged in global projects to develop civil infrastructure involving firms from multiple national institutional backgrounds, for which institutional costs can become highly significant. Also, VDT was extended to model whole enterprises as Project Organization and Workflow for Enterprise Research ("POW-ER") to model highly non-routine work in extremely decentralized "Power to the Edge" organizations (Alberts & Hayes, 2003). This section elaborates the evolution of VDT over the past 20 years, its current status, and ongoing research in this area.

In selecting the kinds of organizations that VDT would initially model, we picked project teams performing routine design or product development work. For this class of organizations, all work is knowledge work so that we could fruitfully use an information-processing abstraction (Galbraith, 1974) of the work. For routine product development, goals and means are both clear and relatively uncontested, so that we could finesse many of the most difficult "organizational chemistry" and "organizational biology" modeling challenges inherent in the kinds of organizations that sociologists have often studied at the enterprise level – e.g., mental health, educational, and governmental organizations. Our intention from the outset was to start with "organizational information flow physics" and then progressively add elements of "organizational chemistry" and "organizational biology" to the modeling framework to extend its applicability to less routine tasks and more dynamic organizations. We have executed several steps of this research vision over the past two decades. Completed and ongoing versions of VDT that progressively addressed additional aspects of task and organizational complexity are shown in Figure 7.

Key Limitations of VDT2/SimVision

The Cohen (1992) and Christiansen (1993) VDT-1,2 framework has been fully validated through all of the steps shown in Figure 6. VDT-2 generates reliable predictions about project work for which: (1) all tasks in the project can be predefined; (2) the organization is static, and all tasks are pre-assigned to actors in the static organization; (3) exceptions to tasks are resolved through the hierarchy and generate extra work volume for the predefined tasks to be carried out by the pre-assigned actors; and (4) actors are assumed to have congruent goals, values, and cultural norms. These conditions fit many kinds of design and product development work. VDT-2 was commercialized as SimVision™ by Vité Corporation through Stanford's Office of Technology Licensing, and it is in use by companies in a variety of industries and governmental organizations including Procter & Gamble, Walt Disney, the US Navy, NASA, and The European Bank for Redevelopment and Construction.

Modeling Moderate Levels of Goal Incongruency

VDT-3 (Thomsen, 1997) extended the range of work processes that could be modeled, to encompass less routine design or product development work, in which tasks are still predefined, but there can be flexibility in how they are executed. Actors can have the same set of goals, but incongruent goal preferences (i.e., a moderate degree of goal incongruency), causing them to disagree about how best to execute tasks in the project plan. Following concepts from economic "Agency Theory", goal incongruency levels between pairs of actors affect both their vertical and horizontal communication patterns.

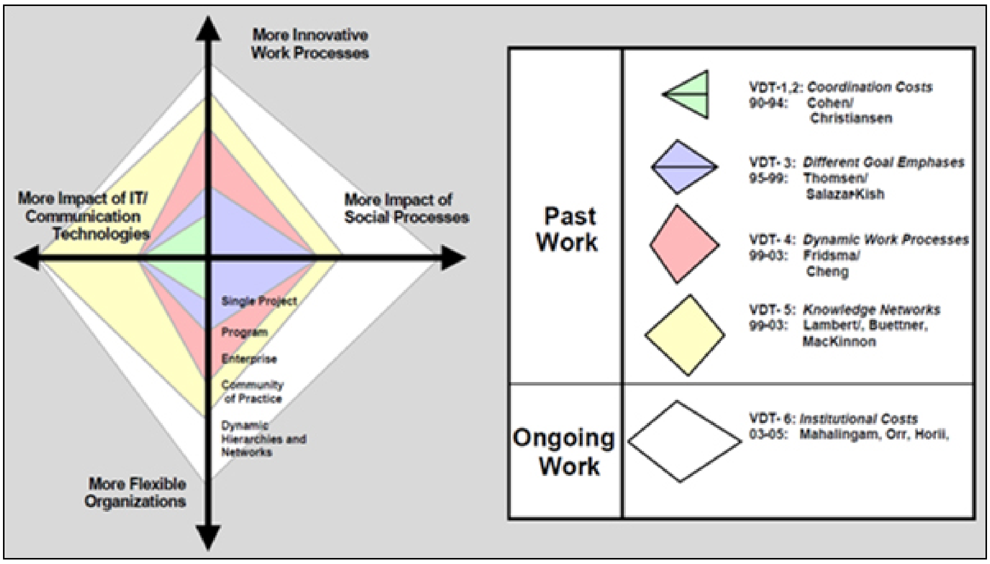

Fig. 7. VDT Research Trajectory

The range of work processes and organizations to which VDT can be applied were expanded step by step: VDT-1,2 for relatively routine, fast-paced project work executed by organizations with hierarchical exception processing, a predetermined and static structure and task assignments, but no significant institutional differences; to VDT 3 for less routine projects where goals of team members might be incongruent; to VDT-4 for non-routine "diagnose and repair" work (e.g., health care delivery or equipment maintenance) executed by more dynamic and adaptable organizations; to VDT-5 in which exceptions can be resolved through team members' knowledge networks rather than just via their supervisors in the hierarchy; to VDT-6 for global projects in which the costs of institutional exceptions arising from the differences in national institutions among team members become significant.

Modeling Less Routine Work Processes: Diagnosis and Repair

A subsequent NSF grant focused on extending the applicability of VDT beyond its previous limits on work-process routineness and static organizational structure. Douglas Fridsma developed VDT-4 to model complex and non-routine health care delivery tasks such as bone marrow transplants and similar complex, multi-specialty, medical protocols. In these work settings, diagnosis tasks indicate needed therapeutic tasks; any unplanned side effects that arise during diagnosis or therapy must be diagnosed and treated contingently. To model the indeterminacy inherent in these kinds of work processes, we had to relax the VDT-1,2,3 constraint that all tasks, actors, and assignments be rigidly pre-specified and remain static. This required several extensions to the VDT-3 framework.

Fridsma (2003) extended the information-processing micro-theory in VDT-3 to include a variety of more complex exceptions that can cause tasks to be added, re-sequenced, deleted, or reassigned, and actors to be dynamically added to the organization and assigned tasks as needed. This extended framework was implemented and internally validated on *toy problems* (see Figure 6). Carol Cheng Cain (Cheng, Cain, & Levitt, 2001) extended Fridsma's work to model context-dependent decision-making —e.g., medical decision-making in intensive care units where organization structure (e.g., level of centralization of decision-making) and staffing (by experienced medical practitioners vs. interns or residents) both change as a function of time of day or day of week— and she *retrospectively validated* VDT-4 predictions against empirical data in several clinical settings (Cheng Cain, 2003).

Modeling Flexible Exception Handling and Knowledge Sharing: "Communities of Practice"

A longer-range goal of our work was to begin modeling even more flexible organizations – dynamically shifting "communities of practice" in which actors can resolve exceptions by communicating not just up the hierarchy, but with anyone from their "knowledge network," either inside or outside their own project organization. Software development teams and some consulting organizations currently approximate this organizational form. Theories based on concepts such as public goods, homophily, or reciprocity can be used to describe how these links form and persist or dissolve in face-to-face working conditions, or in cyberspace for non-co-located teams. We received a NSF Knowledge and Distributed Intelligence (KDI) research grant to work with colleagues from USC, Carnegie Mellon, and the University of Illinois in this exciting new area, and we made significant progress in implementing these extensions. VDT-5, which included these extensions, was reprogrammed and released as **P**roject, **O**rganization and **W**orkflow-**E**xtended **R**esearch (POW-ER) (Ramsey & Levitt, 2005), and has since been used by the US Navy, the US Air Force Research Laboratory, and other governmental organizations.

Modeling Effects of Institutional Differences on Project Team Behavior and Outcomes

Research by Geert Hofstede (1997) and his colleagues provides one clear point of departure for modeling how differences in values and cultural norms can affect the behavior of participants in project teams. Hofstede identified five dimensions of culture that vary systematically between workers from different countries, and which affect individual and team behaviors in global, knowledge-intensive, dynamic, global projects. Hofstede collected large data sets based on IBM employees in more than 50 countries indicating that differences along one or more of these cultural dimensions lead to predictable kinds of misunderstandings, conflict, and loss of motivation in global work teams. This work was subsequently replicated, updated and extended by House, Hanges, Javidan, Dorfman, and Gupta (2004).

Drawing on Hofstede's work and on the results of a series of workshops conducted with Professor Douglass North (a Nobel Laureate in institutional economics at Stanford's Hoover Institute) and Professor Merlin Donald (an eminent Canadian cognitive psychologist) at the Institute for International Studies at Stanford, we developed a set of initial hypotheses about how to model the emergence of "institutional exceptions" and their information-processing costs in global projects within VDT. Scott's (2008) theory of institutions provides a more inclusive conception than Hofstede's limited view of culture as consisting of values and beliefs to explain how sets of mental schemata and individual, group, and legal ideations, norms, and laws drive behavior deemed to be appropriate for persons in different social groups. The doctoral research of Mahalingam (2005) and Orr (2005) found that viewing national differences in terms of institutional differences was far more productive in understanding and predicting cross-national institutional exceptions in projects than viewing them solely through the lens of the Hofstede/House ideas and values constructs.

Our approach was to model institutional work in the same way that we modeled coordination work – i.e., as additional quantities of information to be processed by actors in a project team. Figure 8 shows conceptually how we overlaid institutional work on the production work and coordination work that we had modeled to date. However, in addition to increasing the amount of information to be processed, institutional exceptions may also have the side effect of undermining the motivation of actors who find themselves engaged in continual misunderstandings, conflict, and even sabotage by project team members whose goals, beliefs and values, cultural norms, and legal/regulative systems are significantly different than their own.

Horii (2005) designed and conducted a set of computational experiments in which he modeled US and Japanese institutions (practices and values) and simulated the performance of joint venture teams consisting of US and/or Japanese managers and workers in US- vs. Japanese-style project organizations working on projects with different levels of complexity. His path-breaking work won the best-paper award at CASOS 2005. This line of work has continued since 2005 at the Collaboratory for Research on Global Projects (Scott, Levitt, & Orr, 2011). (See also < http://crgp.stanford.edu >.)

Managers of global projects contending with significant institutional differences need to be realistic about the additional "institutional work" that will be incurred in proceeding with their projects. Forewarned with this kind of prediction, they can set more realistic goals and begin to initiate effective interventions with a clear notion of how long they will take to implement. Additional validation and calibration of Horii's pioneering work will be required for them to do this.

Exploring Fully Automated Organization Design: Developing A Postprocessor for VDT

Organizational design is a complex global and local optimization problem involving continuous and discrete variables. For example, an organizational designer must size functional capabilities, assign staff to tasks, and set communication and control policies. VDT is an analysis tool that can predict schedule cost and process quality performance for a baseline case of an organization and work process, and help to isolate the most severe risks in these three areas. However, VDT cannot suggest how to intervene most productively to change the work process or organization, in order to mitigate any risks that have been identified. The user has to experiment with alternative cases to find better solutions. Searching the solution space manually to find good cases that address schedule, quality, or cost risks for a baseline case is thus a challenging task. It relies on the expertise of the human user and offers no guarantee of optimality or even near-optimality. Because the VDT solution space is so large, and the interaction between its variables is subtle and sometimes counterintuitive, even expert users can fail to discover many potentially superior solutions.

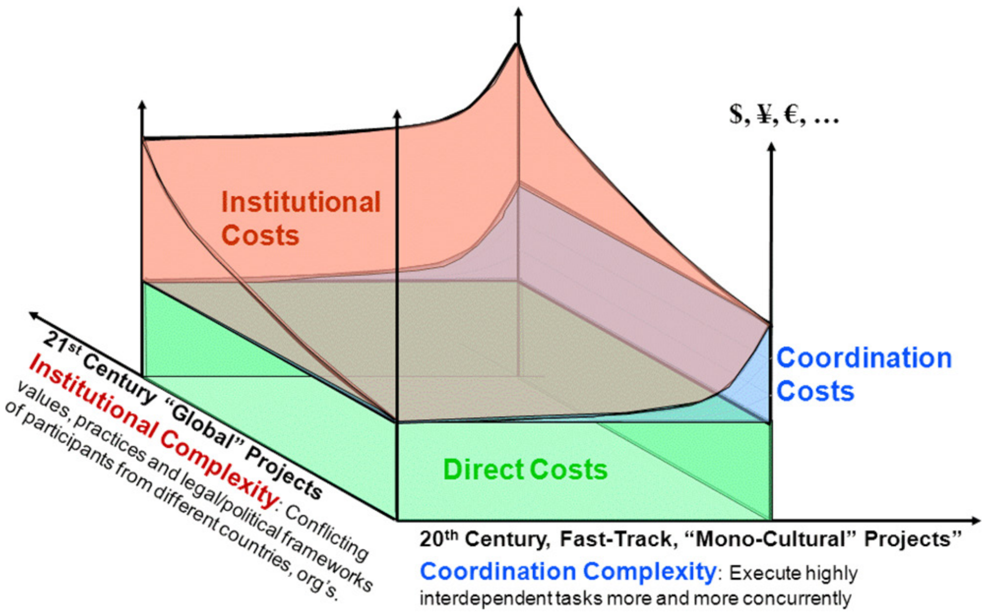

Fig. 8. Costs Arising from Three Kinds of Project Work

Direct costs for projects arise from the cost of assigned actors performing their direct tasks. Additional project costs arise from two kinds of "hidden work": (1) "Coordination Costs" arise from "supervision" – the need for managers to process technical exceptions, and "coordination" – the need for workers and their managers to coordinate interdependencies in highly concurrent project work, and the resulting rework when coordination breaks down; and (2) "Institutional Costs" arise from the need to handle "institutional exceptions" – misunderstandings and conflicts resulting from differences in the national institutions (Scott, 2008) of project team members from different countries, professions, or industry sectors.

During the 1990s, researchers began combining AI and OR techniques to solve several similarly complex kinds of optimization problems (Hooker, 2002). Working in collaboration with Professor John Koza, a pioneer in the development of Genetic Programming, Bijan KHosraviani (KHosraviani & Levitt, 2004; KHosraviani, Levitt, & Koza, 2004) developed a system based on Genetic Programming that was able to evolve VDT models that met a required set of scope, schedule, and cost objectives for a benchmark problem more optimally than multiple teams of human users had been able to do over almost a decade.

Genetic Programming (GP), applied to VDT, attempts to evolve multiple good solutions for a problem via a computational approach that mimics Darwinian evolution of species. In the case of VDT models, GP requires that the user create a "fitness function" that specifies the relative weight to be given to each early completion, low cost, and high quality, as well as to specify any constraints such as the latest possible completion time of the project or maximum number of additional FTEs that could be added to positions. An initial generation of solutions consisting of about 20 different VDT model cases is created as a starting point for the GP. VDT first simulates all of the solution candidates in the initial generation to predict their outcomes. Then the GP evaluates each case's fitness for survival and reproduction as defined by the fitness function, using the outcome predictions for duration, cost, quality, etc. that have been generated by the VDT simulations for the cases in that generation. The cases evaluated as being the fittest by this fitness function preferentially get to propagate themselves to the next generation in one of three ways: they "procreate" – i.e., they exchange genes by combining attributes of the case from a pair of relatively fit "parent" cases into a "child" solution in the next generation; they "mutate" – i.e., one or more randomly selected attributes of a case in a given generation are randomly assigned new values in the next generation; or they can "replicate," in which a relatively fit case is reproduced identically in the next generation so it can continue to pass on its good fitness attributes to future offspring cases. This computational analogy to "evolution of the fittest" proceeds through multiple generations until some cases in the latest generation reach acceptable fitness values.

KHosraviani developed an ingenious dimensional extension to traditional GP, inspired by Professor John Koza's previous work on using GP to "evolve" circuit designs from electronic components (Koza et al., 1996, 1999) akin to evolving multicellular creatures from single cell organisms in a primeval ooze. KHosraviani was then able to apply GP to this multidimensional optimization problem involving both numerical and non-numerical parameters. His GP algorithm was then able to evolve multiple "fit" solutions that surpassed the performance of the best solutions previously identified by multiple student teams in just 20–30 generations. KHosraviani's work was awarded a Silver Medal at the *Genetic and Evolutionary Computation* Conference (GECCO) in 2004.

GP is a computationally intensive process; it required a whole room full of computers linked together as a parallel processor at the time KHosraviani carried out his research. Today GP computations can be carried out on multiple servers "in the cloud."

Modeling Radically Decentralized "Power to the Edge" Organizations

The VDT research has continued since 2005 to develop an extension of VDT called Process, Organization, Work for Edge Research (POW-ER) that could be used to model some of the most decentralized and flexible organizations existing anywhere – so-called "Power to the Edge" organizations such as the US Special Forces team that tracked down and killed Osama bin Laden in 2011 (Alberts & Hayes, 2003) or "Project Management 2.0" organizations (Levitt, 2011) that are increasingly being used to implement "agile software development." POW-ER has now evolved through multiple versions. At the beginning of 2012, Version 3.8 incorporates the ability to model: institutional differences between participants from different nationalities (Horii et al., 2005); learning and forgetting of skills by project team members over the course of an extended project (MacKinnon, 2007); the development of trust between members of a project team who may or may not be co-located (Zolin et al., 2004); and flexible knowledge sharing through networks of human experts and computational support tools such as databases, expert systems, and other computer knowledge archives (Buettner, 2008).

In collaboration with Northrop Grumman Information Technology's Enterprise Applications Group and the US Air Force Research Laboratory we developed an extension of POW-ER to model command-and-control work and other kinds of monitoring and control workflow where predefined sequences of tasks are initiated stochastically by the arrival of intelligence, sensor, or other information, rather than being initiated by the completion of specified predecessor tasks, as in all the previous versions of VDT and POW-ER. We began validating Project Organization and Workflow-Information Driven (**POW-ID**) in the latter part of 2009 (Levitt, Chachere, & Ramsey, 2010).

This overview of the 20-year VDT research program has attempted to explain how a team of researchers was able to begin modeling organizations executing well-specified, complex, but semi-routine project tasks completed by team members with shared goals and institutions, and then to extend the representation and reasoning of the initial theory and tools progressively to address more flexible tasks, more heterogeneous project team membership, and finally more dynamic and decentralized organization structures, as shown in Figure 7. In Hemingway's words, it has been a "movable feast" to participate in this scientific exercise with a remarkable team of faculty and student scholars and collaborators from industry and government.

SUMMARY

This section summarizes where VDT came from, where it has been, its present status, and what might lie ahead for organizational researchers interested in this kind of agent-based simulation.

As we explained in the introduction, VDT arose from the need of managers like Art Smith, facing stringent economic and strategic pressures to execute their large and complex projects more rapidly and concurrently, to find ways to predict the outcomes of proposed organizational cases for their projects and design more effective organizations. Many thousands of projects are planned and executed each year in industries ranging from construction through pharmaceuticals, biotechnology, medical device development, consumer products, computers, software, and other sectors, for which scope is relatively fixed; the structure of tasks, positions, and task assignments are unlikely to change materially over the project duration; and exceptions are processed through one or more hierarchical channels. For this class of projects, VDT-2 has progressed through all the stages of validation in Figure 7 and its commercial descendent, SimVision™, is now routinely being used to support organization design on some of the world's largest projects.

The subsequent versions of VDT, POW-ER, and POW-ID described in the previous section have demonstrated great creativity by multiple PhD researchers in their conceptualization and implementation, and a limited capability to model and simulate more dynamic organizations composed of workers with less homogeneous backgrounds executing less well-structured tasks. However, none of these extensions has been validated extensively enough to support its routine use by practitioners by the end of 2011.

This is not entirely surprising. As stated earlier, semi-routine project organizations lie in the "easy corner" of the space of all organizations, and modeling organizational physics is much easier than modeling organizational chemistry or biology. The frontier of research in the latter two areas is still bounded by top-down rule-based diagnosis of the degree of internal and external fit between attributes of an enterprise's macro-organization structure and its environmental and managerial context, as exemplified by Burton and Obel's (2004) path-breaking Organizational Consultant integration of the contingency theory literature implemented in a book and a software package (OrgCon), and Shenhar's (2001) contingent propositions for designing project organizations and work processes. Looking forward, the article concludes with a set of challenges for researchers interested in advancing the frontier of VDT's model-based style of organization design beyond semi-routine project organizations to help managers like Art Smith design organizations for their increasingly globally networked and fast-changing 21st century projects.

Challenges for Future Research on Organization Design

This section sets out some near-term challenges for future research on model-based organization design that could build on the work described in this article to extend the range of applicability of organizational design theories, methods and tools.

Validating VDT-3 and Subsequent Versions of the VDT and POWER Software

Replication is one of the key means of testing and advancing scientific knowledge in all fields. Replication of early experiments on some of the extensions to VDT-2 described above, and resulting modification and calibration of the representation and reasoning in these simulation modeling tools, can begin to develop sufficient confidence in the predictions of these models of more flexible organizations and dynamic work environments for them to become useful for organization design in settings like healthcare, equipment maintenance, command-and-control organizations, and agile software development.[6]

Modeling Globally Networked Organizational Forms

The world of work and organizations is increasingly global. Moreover, as predicted by Malone, Yates, and Benjamin (1987), computers have driven transaction costs for outsourcing work in many situations toward zero so that today's organizations increasingly deliver their projects using far-flung networks of supply-chain partners rather than just their own direct employees. Agent-based modeling seems ideally suited for modeling the behavior of, and interactions between, global supply-chain partners such as can be found in construction, automobiles, mobile telephones, and many other kinds of mature products assembled from relatively standardized components. This represents an exciting area of near-term application for agent-based simulation technology (Chinowsky & Taylor, 2012). Secondly, networked organizations in mature industries face significant challenges when attempting to innovate systemically rather than at the module level (Sheffer, 2011). Again, agent-based models of project networks can shed further light on this important subject.

Dynamically Predicting and Controlling Project Organizations

Autopilots used to help pilots or captains guide and control airplanes or ships combine real-time data from a variety of sensors and other data sources about the airplane's internal operation and external variables (e.g., current engine and control surface settings, and en route traffic congestion or meteorological conditions) with the ability to predict the impact of changes in engine power, control surface orientation, etc. on the vehicle's trajectory, arrival time, fuel supply, etc., and to issue alerts to the pilots and or ground controllers when out-of-bounds conditions arise. Similarly, it would be worthwhile attempting to link tools like VDT (or, more likely, its commercial SimVision implementation) to the parent organization's "sensor network" and data – its IT systems for enterprise resources planning, customer relationship management, human resources, and the like – to help managers control their organizations dynamically in real time in accordance with both the organization's top-level strategic objectives and each project's objectives and constraints along with its actual progress to date in meeting those objectives and constraints.

Nissen and Burton (2011) have developed the concept of "dynamic fit" using control of the trajectory and orientation of an airplane as an analogy. They use the notion of "opportunity costs" as a kind of overarching fitness function for operationalizing organizational tradeoffs over time. Future versions of VDT could incorporate this notion in guiding managers' interventions toward more optimal organizational configurations.

6 The author will gladly make current versions of the POWER software, implemented in the Python language, available to researchers interested in pursuing ongoing validation and extension of these agent-based modeling and simulation frameworks.

Developing Next Generation Simulation Models

In 1988, SmallTalk (Goldberg & Robson, 1983) and IntelliCorp's Knowledge Engineering Environment (KEE), which was implemented in LISP, were just about the only object-oriented computing languages available to our team, and the only object-oriented simulation language that could reason about non-numerical variables was IntelliCorp's KEE-SimKit™. KEE-SimKit provided us with a powerful prototyping language for Cohen's (1992) prototype of VDT-1, but it ran only on expensive and custom LISP-processing hardware from Xerox or Symbolics, and the simulations executed painfully slowly. This was a problem even for researchers, because the stochastic nature of VDT required us to run at least 100 simulations of each model case and develop average and standard deviation measures to interpret the results with any statistical reliability. When SimVision was commercialized in 1996, it was developed in C++, the object-oriented language based on C that has become widely used since the mid-1990s. This required the agent-based simulation functionality to be developed essentially from scratch; the advantage was that simulations implemented and compiled in C++ executed rapidly enough to be useful not only to researchers but also to managers.

If the VDT team were starting work today, we would be faced with a plethora of object-oriented programming environments that can be executed rapidly on desktop, laptop or "in-the-cloud" computers, and even multiple agent-based simulation environments such as the Santa Fe Institute's SWARM language for linked, multilevel, agent-based simulations (Minar, Burkhart, Langton, & Askenazi, 1996). The graphical tools for building model canvases and displaying simulation outputs have also evolved dramatically. Early versions of SimVision used Microsoft's Visio™ for this. Later versions deployed a custom-developed user interface built on graphical libraries from open-source or commercial developers. Researchers interested in developing models of supply-chain networks (Chinowsky & Taylor, 2012), knowledge networks, or other networks can similarly access powerful off-the-shelf social network modeling, analysis, and visualization tools such as UCINET (Borgatti, Everett, & Freeman, 2002). So progress in this field has the potential to accelerate dramatically.

CONCLUSION

As described above, the "information flow physics" of project organizations are now relatively well understood and modeled. The author's hope and strongly held belief is that – like their natural science counterparts – "organizational chemistry" (goal conflict, institutional differences, and the like) and "organizational biology" (individual learning, organizational learning, evolution and regeneration of networks of organizations) will eventually yield to robust and accurate enough agent-based modeling, analysis, and validation that simulation of these phenomena will become useful to managers like Art Smith in designing their globally networked 21st-century project organizations. There is much exciting work to be done!

REFERENCES

Alberts D, Hayes R. 2003. *Power to the Edge: Command… Control… in the Information Age*. Center for Advanced Concepts and Technology, Command and Control Research Program, US Department of Defense. http://www.dodccrp.org/files/Alberts_Power.pdf

Programming to Scheduling Problems. Kluwer Academic Publishers, Norwell, MA.

Borgatti SP, Everett MG, Freeman LC. 2002. *UCINET for Windows: Software for Social Network Analysis*. Analytic Technologies, Harvard, MA.

Buettner RR. 2008. The impact of communal knowledge mechanisms on project work (Doctoral dissertation, Department of Civil and Environmental Engineering, Stanford University.)

Burton RM, Obel B. 1995. The validity of computational models in organization science: from model realism to purpose of the model. *Computational and Mathematical Organization Theory* 1(1): 57–71.

Burton RM, Obel B. 2004. *Strategic Organizational Diagnosis and Design: The Dynamics of Fit, Third edition, (with OrgConSoftware)*. Kluwer Academic Publishers, Boston, MA.

Burton RM, Obel B. 2011. What-Is, what-might-be, what-should-be studies – and triangulation? *Organization Science* 22(5): 1195–1202.

Cardinal LB, Turner SF, Fern MJ, Burton RM. 2011. Organizing for product development across technological environments. *Organization Science* 22(4): 1000–1025.

Carroll T, Burton RM. 2000. Exploring 'complex' organizational designs. *Computational and Mathematical Organization Theory* 6(4): 319–337.

Carroll TN, Gormley TJ, Bilardo VJ, Burton RM, Woodman KL. 2006. Designing a new organization at NASA: an organization design process using simulation. *Organization Science – Special Issue on Organizational Design* 17(2): 202–214.

Carroll TN, Burton RM. 2012. A contingency approach to designing project organizations: theory and tools. *Engineering Project Organization Journal* 2(1-2): 5–14.

Cheng CHF, Levitt RE. 2001. Contextually changing behavior in medical organizations. *Proceedings of the 2001 Annual Symposium of the American Medical Informatics Association*, Washington, DC.

Cheng Cain CF. 2003. Representation and reasoning about contextually changing organizational behavior using simulation models of medical work (Doctoral Dissertation, Department of Biomedical Informatics, Stanford University.)

Chinowsky P, Taylor JE. 2012. Networks in engineering: an emerging approach to project organization studies. *Engineering Project Organization Journal* 2(1-2): 15–26.

Christiansen TR. 1993. Modeling efficiency and effectiveness of coordination in engineering design teams (Doctoral Dissertation, Department of Civil & Environmental Engineering, Stanford University.)

Cohen GP. 1992. The Virtual Design Team: an information processing model of the design team management (Doctoral Dissertation, Department of Civil & Environmental Engineering, Stanford University.)

Cohen KJ, Cyert RM. 1965. Simulation of organizational behavior. In J. G. March (Ed.), *Handbook of Organizations*. Rand McNally, Chicago, IL, 305–334.

Cohen MD, March JG, Olsen JP. 1972. A garbage can model of organizational choice. *Administrative Science Quarterly* 17 (1): 1–25.

Epstein JM, Axtell R. 1996. *Growing Artificial Societies: Social Science from the Bottom up*. Brookings Institution, Washington, DC.

Frisma D. 2003. Organizational simulation of medical work: an information-processing approach (Doctoral Dissertation, Department of Biomedical Informatics, Stanford University.)

Galbraith JR. 1974. Organizational design: an information processing view. *Interfaces* 4: 28–36.

Goldberg A, Robson D. 1983. *Smalltalk-80: The Language and its Implementation*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA.

Hofstede G. 1997. *Cultures and Organizations: Software of the Mind*. McGraw Hill, New York.

Hooker J. 2002. Logic, optimization, and constraint programming. *INFORMS Journal on Computing* 14(4): 295–321.

Horii T. 2005. Impact of Multiple Normative Systems on the Organizational Performance of International Joint Ventures (Doctoral Dissertation, Department of Civil & Environmental Engineering, Stanford University.)

Horii T, Jin Y, Levitt RE. 2005. Modeling and analyzing cultural influences on project team performance. *Computational and Mathematical Organization Theory* 10 (4): 305–321.

House RJ, Hanges PJ, Javidan M, Dorfman PW, Gupta V. 2004. *Leadership and Organizations – the GLOBE study of 62 Societies*. Sage Publications, Inc.

Ibrahim R, Nissen ME. 2007. Discontinuity in organizations: developing a knowledge based organizational performance model for discontinuous membership. *International Journal of Knowledge Management* 3(1): 18–36.

Jensen KW, Håkonsson DD, Burton RM, Obel B. 2010. The effect of virtuality on the functioning of centralized versus decentralized structures – an information processing perspective. *Computational and Mathematical Organization Theory* 16(2): 144–170.

Jin, Y, Levitt RE. 1996. The Virtual Design Team: A Computational Model of Project Organizations. *Journal of Computational and Mathematical Organization Theory* 2 (3):

171–195.

KHosraviani B, Levitt RE. 2004. Organization design using genetic programming. Proceedings of North American Association for Computational Social and Organizational Science (NAACSOS) Conference, Pittsburgh, PA, June 27–29.

KHosraviani B, Levitt RE, Koza JR. 2004. Organization design optimization using genetic programming. In M. Keijzer (Ed.), *Late-Breaking Papers at the 2004 Genetic and Evolutionary Computation Conference (GECCO-2004)*. Seattle, WA International Society of Genetic and Evolutionary Computation.

Kim J, Burton RM. 2002. The effect of task uncertainty and decentralization on project team performance. *Computational and Mathematical Organization Theory* 8 (4): 365–384.

Klein R. 2000. *Scheduling of Resource-Constrained Projects*. Kluwer Academic Publishers, Norwell, MA.

Koza JR, Bennett FH III, Andre D, Keane MA, Brave S. 1996. *Genetic Programming III Videotape: Human-Competitive Machine Intelligence*. Morgan Kaufmann, San Francisco.

Koza JR, Bennett FH III, Andre D, Keane MA. 1999. *Genetic Programming III: Darwinian Invention and Problem Solving*. Morgan Kaufmann, San Francisco.

Kunz JC, Christiansen TR, Cohen GP, Jin Y, Levitt RE. 1998. The Virtual Design Team: a computational simulation model of project organizations. *Communications of the Association for Computing Machinery (CACM)* 41(11): 84–91.

Levine SS, Prietula MJ. 2011. How Knowledge Transfer Impacts Performance: a multi-level model of benefits and liabilities. Organization Science, *Articles in Advance*: 1-19.

Levitt RE, Cohen GP, Kunz JC, Nass CI, Christiansen T, Jin Y. 1994. The 'Virtual Design Team': simulating how organization structure and information processing tools affect team performance. In K. M. Carley and M. J. Prietula (Eds.), *Computational Organization Theory*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Levitt RE, Thomsen J, Christiansen TR, Kunz JC, Jin Y, Nass C. 1999. Simulating project work processes and organizations: toward a micro-contingency theory of organizational design. *Management Science* 45 (11): 1479–1495.

Levitt RE, Fyall M, Bjornsson P, Hewlett W III, Casebolt T. 2002. When information flow in project organizations becomes turbulent: toward an organizational Reynolds Number. *Proceedings of the Conference on Agent-Based-Modeling in the Social Sciences*, UCLA: 3–10.

Levitt RE, Chachere J, Ramsey M. 2010. Organizational research generator for simulation in the military. *Collaboratory for Research on Global Projects, Stanford University, Working Paper #0060*. http://crgp.stanford.edu/publications/working_papers/Chachere_ORGSIM_WP0060.pdf

Levitt RE. 2011. Toward project management 2.0. *Engineering Project Organization Journal* 1(3): 197–210.

Long CP, Burton RM, Cardinal LB. 2002. Three controls are better than one: a computational model of complex control systems. *Computational and Mathematical Organization Theory* 8(3): 197–220.

MacKinnon DJ. 2007. How individual skill growth and decay affect the performance of project organizations (Doctoral dissertation, Department of Civil & Environmental Engineering, Stanford University.)

Mahalingam A. 2005. Understanding and mitigating institutional costs on global projects (Doctoral dissertation, Department of Civil & Environmental Engineering, Stanford University.)

Malone TM, Yates J, Benjamin RI. 1987. Electronic markets and electronic hierarchies. *Communincations of the Association of Computing Machinery* 30 (6): 484–497.

March JG, Simon HA. 1958. *Organizations*. Wiley, New York.

Masuch M, LaPotin P. 1989. Beyond garbage cans: an AI model of organizational choice. *Administrative Science Quarterly*, 34 (1): 38–67.

Minar N, Burkhart R, Langton C, Askenazi M. 1996. The Swarm simulation system: a toolkit for building multi-agent simulations. Santa Fe Institute, Santa Fe, NM. www.swarm.org/images/b/bb/MinarEtAl96.pdf

Nissen ME, Levitt RE. 2004. Agent-based modeling of knowledge-flows: Illustration from the domain of information systems design. *Proceedings of Hawaii International Conference on System Sciences*, Waikoloa, HI.

Nissen M, Burton RM. 2011. Designing organizations for dynamic fit: system stability, maneuverability, and opportunity loss. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans* 41(3): 418–433.

North DC. 1990. *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, Cambridge, UK.

Orr RJ. 2005. Unforeseen conditions and costs on global projects: learning to cope with unfamiliar institutions, embeddedness and emergent uncertainty (Doctoral dissertation, Department of Civil & Environmental Engineering, Stanford University.)

Ramsey M, Levitt RE. 2005. A computational framework for experimentation with edge organizations. *Proceedings of the International Command & Control Research Symposium*, McLean, Virginia.

Schrage M. 2000. *Serious Play: How the World's Best Companies Simulate to Innovate*. Harvard Business School Press, Boston, MA.

Scott WR. 2008. *Institutions and Organizations,* 3rd Ed. Sage Publications.

Scott WR, Levitt RE, Orr RJ. (Eds.) 2011. *Global Projects: Institutional and Political Challenges*. Cambridge University Press, Cambridge, UK.

Sheffer DA. 2011. Innovation in modular industries: implementing energy-efficient innovations in US buildings (Doctoral dissertation, Department of Civil & Environmental Engineering, Stanford University.)

Shenhar A. 2001. One size does not fit all projects: exploring classical contingency domains. *Management Science* 47(3): 394–414.

Simon HA. 1947. A*dministrative Behavior: A Study of Decision-making Processes in Administrative Organization*. Macmillan, New York.

Steward DV. 1981. The design structure system: a method for managing the design of complex systems. *IEEE Transactions on Engineering Management* 78(3), 71–74.

Thompson JD. 1967. *Organizatons in Action*. McGraw-Hill, New York.

Thomsen J. 1998. The virtual team alliance (VTA): modeling the effects of goal incongruency in semi-routine, fast-paced project organizations (Doctoral dissertation, Department of Civil & Environmental Engineering, Stanford University.)

Thomsen J, Kwon Y, Kunz JC, Levitt RE. 1997. Simulating the effects of goal incongruency on project team performance. Fourth Congress on Computing in Civil Engineering, ASCE, Washington, DC, 643–650.

Thomsen J, Levitt RE, Kunz JC, Nass CI, Fridsma DB. 1999. A trajectory for validating computational emulation models of organizations. *Journal of Computational and Mathematical Organization Theory* 5(4): 385–401.

Wong S, Burton RM. 2000. Virtual teams: what are their characteristics and impact on team performance? *Computational and Mathematical Organization Theory* 6(4): 339–360.

Zolin R, Hinds PJ, Fruchter R, Levitt RE. 2004. Interpersonal trust in cross-functional, geographically distributed work: a longitudinal study. *Information and Organization* 14(1): 1–24.

RAYMOND E. LEVITT

Kumagai Professor of Engineering
Stanford University
E-mail: ray.levitt@stanford.edu

DESIGN OF INDUSTRIAL AND SUPRA-FIRM ARCHITECTURES

GROWTH AND SUSTAINABILITY

JOHN A. MATHEWS

Abstract: The scope of organization design has expanded steadily from work-flow issues and job specifications to firm-level considerations and now to supra-firm industrial structures, where such issues as modularity and clustering loom large. Economic analysis has made little headway in analyzing how increasing returns may be generated through supra-firm structures such as networks and clusters, nor in the question of how their industrial architecture (modular vs. integral, open vs. closed) affects economic performance. The focus here is on the supra-firm industrial architectures that have arisen, either spontaneously through the evolution of capitalism or through purposeful design, involving both state and private actors. Striking cases such as the Chinese automotive industry, which started with the production of conventional automobiles and motorcycles and now encompasses both two-wheeled and four-wheeled electric vehicles, provide testimony to the power of some industrial configurations to outperform others. My analyses and arguments are placed in the global context of the urgent need to find ways to accelerate the uptake of green technologies (such as electric vehicles) in order to reduce emissions of greenhouse gases and at the same time promote the industrialization of countries still at lower levels of income and wealth.

Keywords: Industry architecture; organization design; modularity; industrial clusters; integrality; Chinese automotive industry; electric vehicles; electric two-wheeled vehicles

Use of electric two-wheeled vehicles is rising rapidly in China, making them the world's most significant form of electric-powered transport. The success of electric two-wheelers (E2W) is expected to have a big impact in driving down the cost of batteries and creating a ready market for four-wheeled vehicles, which is now getting under way in serious fashion in China. The growth in two-wheelers is phenomenal, even by the extraordinary standards of growth in China, rising from around 2 million units in 2002 to 14 million just four years later, with production and sales reaching 30 million vehicles by 2010. This is market expansion that calls for explanation beyond that of low costs.

There are several factors behind this market surge, among which the banning of conventional motorcycles in several municipal areas in China (on account of their noise and pollution) is significant. But the most important driving factor is the *modular industry structure* that has developed, allowing hundreds of new entrants to flock to the industry and assemble e-bicycles from modular components and standardized interfaces. These components, which constitute a few core platforms from which the e-bikes are created, are themselves adaptations of platforms developed by foreign firms. The Chinese E2W industry is a demonstration of the power of industry architecture, in this case a multi-level, quasi-open, modular structure where new firms can enter the industry relatively easily and produce new models at low cost, thereby accelerating the diffusion of new, green transport solutions (Weinert, Ogden, & Burke, 2008). The result is that China, which now has the world's highest carbon emissions from its intensive growth in fossil fuel usage, could become the world's cleanest transport sector (heavily involving private electric vehicles and public fast train systems), leapfrogging the rest of the world to this status through the aid of modular industry

JOD, 1(2): 42-63 (2012)
DOI: 10.7146/jod.6454

architecture. Environmentally speaking, this is good for China and for the world.

Time was when "organization design" referred to the design of the entities that populate the international business system – the firms and institutions making up the modern global economy. Organizational analysis emerged at the same time as the discipline of management appeared, coincident with the rise of large-scale production and marketing firms that transformed first the U.S. economy and then the European and Japanese economies, and has since spread worldwide (Fligstein, 1985; Franko, 1976). The concept and literature of organizational design emerged in the 1970s, with the realization that there were different approaches to capturing managerial and organizational efficiencies, even in what looked like standard-issue mass production firms. As "lean production," "mass customization," and "total quality management" came of age, along with notions of sociotechnical foundations of organization, so the prospects for organizational design broadened and brightened. Then the links with the wider economy were emphasized, as "industrial marketing and purchasing" and "supply chain management" emerged, so that the emphasis of organization design shifted from the firm itself to the organization within its network of inter-firm linkages.

Organization design still refers largely to possibilities inherent in the design and redesign of single firms or organizations. But there is now the inescapable added dimension of the organizational setting – or what could be termed, at a higher level of recursion, the issue of *supra-firm design*. It is this level that I address in this article.

CONCEPT OF SUPRA-FIRM DESIGN

The impetus for this development lies partly in the fact that some supra-firm designs clearly work better than others. Take the case of innovation networks, which could be manifested, for example, as R&D consortia. Here the issue of supra-firm design concerns the links between the collaborating firms, in terms of their R&D activities (where they might collaborate while remaining sturdy competitors in terms of products), and the design of these inter-firm linkages or routines. In Taiwan, for example, there was a rapid evolution of supra-organizational design of the country's R&D consortia that constituted an important institutional aspect of the government's "catch-up" strategy (Mathews, 2002a, 2003). Taiwan's consortia evolved from informal networks with minimal commitment to well-defined networks with substantial financial and technological commitment, from consortia with vague goals to organizational arrangements with precise goals. Improvements in the design of these supra-firm structures could be termed "economic learning" in contrast to organizational learning. The Taiwan R&D consortia provide an excellent example of design at the supra-firm level, where the emphasis is on collective improvement or learning. As such, it is central to national economic and industrial success.

Another case is provided by supply chain networks, where some companies have been able to generate enormous advantages by their ability to design a particular configuration of capabilities and activities along a supply chain and choosing where to insert themselves in such a configuration. Fine (2000, 2005) employs the term "supply chain design" to capture such strategic choices, where firms can draw advantages by localizing specialized capabilities in designated suppliers while emphasizing their own capabilities in branding and design. Here the processes of intermediation and disintermediation can follow each other sequentially.[1]

The impetus for supra-firm design as an analytic category in its own right derives also from an even more compelling development, namely that the current architecture of capitalist industry is non-sustainable (see Stern (2007) for a summary of the scientific evidence). It is leading to an over-reliance on fossil fuels, resulting in excessive combustion and emissions of greenhouse gases and to an over-reliance on a linear economic model where we take resources at one end from something called "nature" and dump wastes at the other end, again in "nature" – and never make the connection that these processes might be linked at a higher level of recursion. The burning of fossil fuels is leading to global warming or, more simply, to

1 Fine contrasts modular with closed product architectures and their linked supply chain organizational architectures. Voordijk, Meijboom, and De Haan (2006) provide an interesting test of the proposition that modular supply chains enhance the effect of modular product architectures, in the case of the construction industry, Mathews (1996) treats the same issues utilizing the terminology "holonic."

the fact that we are "cooking" the planet (Krugman, 2010).[2] The degradation and throwaway of resources cannot be allowed to continue if we want civilization to survive.

Concern over these issues started at the margins, in what was called the "environmental movement," but it has moved to the mainstream, and now with the rise of China and India, to the very center of concern. The environmental repercussions of conventional organization design (burning the fuels and wasting the resources) could be tolerated while they were practiced by a minority of the earth's inhabitants, but now that the Chinese and Indians have started to demand their share of industrial riches and have moved decisively to do so, then it seems that the design of fossil-fuelled industrial structures was not so effective after all. It cannot scale to a population of seven billion. Herein lays the extremely "inconvenient" truth facing industrial civilization (Gore, 2007).

Attention to supra-firm design – design at the industry level and higher – in such a way that diffusion of new low-carbon technologies might be accelerated is now emerging as a critical and indeed existential issue. Concern with the world of business that lies outside the firm has conventionally been the domain of economics – whether in its "micro" version at the level of individual actors or its "macro" version at the level of aggregates such as national income, taxes, and interest rates. But economics utilizes an outmoded framework of analysis that is mechanistic and equilibrium-based at the expense of consideration of the evolutionary dynamics and strategic interactions of firms within the real economy (Mathews, 2006a,b; 2010). The framework of conventional economics, with its linear thinking and refusal to take the ecological setting of industrial activity seriously (marginalizing it as externalities), is now an obstacle to redesigning the industries that have created the problem. Moreover, industrial economics in its conventional form is concerned with firms and markets and the price-guided competition they generate. But today the most interesting processes occur in the economic space *between* firms and markets – in networks, clusters, platforms, circular economy loops, and other supra-firm structures that are studied in strategic management but not (yet) in mainstream economics.

Moving forward, a new paradigm is called for, and one that suggests itself is the paradigm that has guided life through its millions of years of evolving more resilient and adaptable forms. This paradigm is the imitation of life and its processes, or *biomimesis*. Scholars such as Reap, Baumeister, and Bras (2005) outline an approach to the design of the industrial system as a whole, along biomimetic lines, in what they call "holistic biomimesis" – an approach that moves beyond the design of individual products and processes to consider the systemic and ecosystemic levels that are normally left out of industrial analysis and practice. They identify seven characteristics of ecosystems, drawing on the work of Benyus (1997), as formulating "conditions conducive to life." These conditions have the property that they have been tested by several billion years of evolutionary development – in other words, they are resilient.

In this article, I consider aspects of industry-level redesign of organizational architectures – modular-open architectures, industrial clusters, and eco-connections between firms – and demonstrate that each carries an overtone of biomimesis. In this way, industry design could help drive the transformation of industries that is needed if newly aspiring industrial powers like China and India are to accomplish their goals – most importantly, without degrading the planet.

REDESIGN OF INDUSTRIES AND ACCELERATION OF THE SHIFT TO GREEN TECHNOLOGIES

Industrial capitalism has revealed itself to be the most powerful transformative agent found in the world today. Its appearance in Britain in the second half of the 17th century, powered by access to fossil fuels, unleashed astonishing gains in productivity associated with rises in income, so that it was widely emulated. Polanyi (1944, 1957, 2001) aptly called this *The Great Transformation*, in the sense that nothing would be the same again. Capitalism was

2 See Paul Krugman, Who cooked the planet?, *New York Times*, July 25, 2010; available at: http://www.nytimes.com/2010/07/26/opinion/26krugman.html

indeed an amazing invention of humankind. Its appearance in cities led to demands for independence and liberties that today we take for granted in the West and which are now spreading worldwide. It ushered in the Industrial Revolution, which applied fossil fuels to production, along with new mechanical inventions, thus starting the world on a trajectory of industrialization and modernization that is bringing more and more of the world's people into its orbit.

The process of industrialization lifted close to one billion people in Western Europe, North America, and Japan out of the "Malthusian trap" that pinned income to population, and set them on a trajectory of rising per capita wealth. This created a "Great Divergence" between the West and "the Rest," accounting for the extreme disparities in wealth, income, and power that have characterized the modern world. In the 20th century, while serious efforts were made to industrialize in many parts of the world, it was only in East Asia that catch-up, or convergence, was achieved. Now in the 21st century these efforts have spread to China and India, and a "Great Convergence" is under way, reversing the trajectories of the past 200 years (Pomeranz, 2000; Wolf, 2011).[3]

However, if up to six billion people are to be raised to middle-income status by 2050, as envisaged by economists such as Spence (2011), then the model of industrialization has to scale sixfold. The industrial system fashioned over the past two centuries is now being stretched to accommodate the rise of new industrial powers such as China, India, and Brazil, and more of the emerging market countries after them. While the impact of the industrial system on its ecological setting – the environment – could be more or less ignored in the early phase of industrial expansion, now that it is filling the earth this is no longer a feasible option. The Western model cannot scale to accommodate the aspirations of China, India, and all the other peoples waiting in line for their turn to enjoy the fruits of industrialization. The urban congestion, pollution, waste generation, demands on fossil fuels, and the resource wars that would have to be fought to extend and defend oil supply lines are the consequences of extending the business-as-usual pathway and the reasons why the Western model cannot scale.

Rather than take the usual policy-oriented approach to these issues, I propose to examine the problem from the fresh vantage point of design – from the perspective of industrial and organizational architectures. Industries have grown within the setting of the capitalist mass production system with few constraints. Corporations are protected by limited liability, and on that basis they feel free to take resources as they wish, which in effect means plundering the earth's natural capital both as a source and a sink. This is one level of concern, which has to be addressed by the redesign of industries along "circular economy" lines where outputs from one process are fed as inputs into another process, in emulation of biological cycles (Mathews & Tan, 2011).

Another level of concern is raised by the need to identify the green technologies that can counter the tendencies to excessive resource throughput and carbon emissions, and to accelerate their uptake, in China and other countries where such technologies are most needed. This level of concern can be addressed through the redesign of industries along open-modular or quasi-open-modular lines, as opposed to the traditional closed-integral organizational architectures that dominated in the early years of industrial capitalism. There is now abundant evidence that diffusion of technologies is accelerated by open-modular industry architectures, as evidenced by the rapid uptake of new industries such as personal computers, video players, and cell phones by Taiwan, Korea, and now China, based on the modularization of these sectors. The same phenomenon is evident in new energy vehicles in China, including four-wheeled and two-wheeled electric vehicles, which promise to have an enormous impact in reducing that country's carbon emissions as well as urban pollution.

Wider concerns with social and economic inequalities have become a major destabilizing influence as industrial capitalism has globalized and brought new populations within sight of a middle-class income, provided their countries can master a new green development model that is not as resource-intensive as the earlier Western model that will not scale to accommodate a world of eight or nine billion aspirants. The architecture of industries based

3 See Martin Wolf, In the grip of a Great Convergence, Financial Times,January 4, 2011; available at: http://www.ft.com/cms/s/0/072c87e6-1841-11e0-88c9-00144feab49a.html#axzz1yx8OrRrZ

on isolated firms working to an anonymous market, along the model utilized in industrial economics, is giving way in these emerging countries to a more systematically pursued industrial clustering architecture, pioneered in China by Special Economic Zones (SEZ) and now diffusing to India and beyond. Within these SEZs, cities and towns are able to promote industrial clusters, on the model of the industrial districts pioneered in Europe and the United States in the 19th century, enabling firms to generate increasing returns as they multiply their interconnections and enable specialized firms to emerge as markets expand. There is again a clear biomimetic aspect to this, in the sense of the clustering of nerve cells to create ganglia in which resides intelligence – or, in the business analogue, "value-added."

MODULARITY AND ITS IMPACT ON TECHNOLOGY DIFFUSION

Modular systems have been with us for a long time. Design platforms emerged first in the military (which has driven the pace of technological change in multiple sectors) and then spread to the automotive sector, where modular platforms enabled companies to offer wide consumer choice based on a few core modules for the chassis, transmission, and engine. Modular design came to particular notice with the rise of consumer products like stereo sound systems and the personal computer with IBM-compatibility.[4] Firms like Dell were able to build a new business model out of the design of systems involving modularity, while in the semiconductor sector new IC foundry firms like TSMC, UMC, and Chartered, from Taiwan and Singapore, pioneered the strategy of *silicon modularity*, where chips are designed through the use and re-use of in-silicon system components, each one of which is IP-protected. Modularity and platforms emerged together, with each driving the other in a process now described in the business literature as co-evolution. In the canonical description given by Baldwin and Clark (2000), design rules for modularity encompass three categories: (1) *modular architecture*, which specifies what the modules do and how they fit together; (2) *interfaces*, which specify the rules of modular interconnection; and (3) *standards*, which test and prove a module's ability to fit within the overall system and enable one module to be compared with another within the system context. It is through the judicious understanding of the workings of these design rules that firms and systems designers are able to capture advantages from modularity not available in non-networked, integrated products. Gawer and Cusumano (2002, 2008) took these arguments further in their description of the emergence of platforms, both in their closed and open variants, as creating a new source of competitive advantage shared by many firms. Platforms, such as the Wintel platform in PCs, or the Apple platform that links the iPad, iPhone, and iPod, are in fact ubiquitous. The more an industry moves toward modularity, the greater is the likelihood that one firm will be able to initiate a new level of integration to bring components together into a platform. And as platforms emerge, the more pressure they place on the industry to modularize.

Simon (1962) provided the first modern discussion of these systems, and at the same time offered a plausible evolutionary account of their emergence in his famous fable of the two watchmakers. One watchmaker, Tempus, builds watches out of their constituent elements, while the other, Hora, builds his watches out of modular components containing ten elements each. Simon then discusses the relative difficulties the watchmakers face, particularly when encountering interruptions that disrupt the flow of work. Hora pulls ahead of Tempus rapidly because interruptions force Tempus to start each watch over again. Simon draws intriguing analogies from this case for the field of biological evolution, arguing that complex organisms arise precisely because of their modular, or cellular, structure.

4 See Langlois and Robertson (1995) for a general description of modular systems and application to the stereo and computer industries, while Chesbrough and Kusunoki (2001), Brusoni and Prencipe (2001), Langlois (2002, 2003), Schilling (2000), Chesbrough (2005), Baldwin and Woodard (2007, 2009), Campagnolo and Camuffo (2010), and Chesbrough and Prencipe (2008) provide overviews of the strategic and management issues involved. These concepts have been captured under a variety of names: Mathews (1996) termed them cases of "holonic" organizational architectures, while Miles, Snow, Mathews, Miles, and Coleman (1997) termed them "cellular" systems. All of these terms are getting at the same point, namely the building of systems out of re-usable and substitutable components that are themselves also systems, with their own systemic properties. Garud and Kumaraswamy (1995) go further and offer the beginnings of a theory as to why such systems offer advantages, introducing the concept of "economies of substitution."

The issue of industry design is brought out once we recognize that different organizational architectures are possible, and these lead to different business and economic outcomes. We are talking about inter-firm organizational architectures – the relations that firms have with each other, mediated through the character of the products they are engaged in producing. The key distinction to be made is between modular and integral architectures on the one hand, and open and closed architectures on the other. While there is no necessity for a modular product to be produced by modular industry architecture, there is obviously an "elective affinity" between the two approaches.[5] Japanese scholar Takahiro Fujimoto (1999, 2006) has done a service in clarifying the main issues. In the modular architecture, when talking of a product, there is a one-to-one correspondence between functional and structural elements. In the personal computer, for example, the function of processing is handled by the CPU, the function of input by the keyboard, the function of display by a flat panel display, and so on. In the integral architecture, there is many-to-many correspondence. In the automotive vehicle, for example, fuel efficiency is linked to the engine but also to the design of the body and suspension, while the suspension affects not just fuel efficiency but also the quality of the ride and the handling of the vehicle. On the other hand, there are open and closed systems or architectures. The Apple Macintosh worked with a closed system while the IBM PC moved toward an open system, based on the fact that IBM believed it was to its advantage to open the system up to components suppliers working to IBM standards. Fujimoto (1999, 2006) provides a threefold classification of product architectures arrayed in a 2x2 matrix, as shown in Figure 1.

	Integral	**Modular**
Closed	small cars motorcycle game software compact consumer electronics	mainframe computer machine tools LEGO (building-block toy)
Open		personal computer (PC) bicycle PC software internet

Fig. 1. Three basic types of product architecture: closed-integral, closed-modular, open-modular.
Source: Fujimoto (1999, 2006)

The three basic types of product architecture have a strong bearing on the organizational architecture of the industries producing such products. Industries generally start by producing products in closed-integral form (with a large company supplying the integral product, serviced by a supply chain in which there may be extensive vertical integration). Think of computers produced by IBM or PCs produced by Apple in the 1980s. Such architecture may evolve into a closed-modular form, as the lead firm seeks to economize in its own production through simplifying its product architecture, with a view to outsourcing the production of some of the lower-cost modular components. Think of IBM producing its System 360 computer along these lines, and then its IBM PC in the early 1980s. Such an architectural innovation, to use the terminology of Henderson and Clark (1990), frequently evolves into an open-modular

5 See, for example, Henderson and Clark (1990) and Sanchez and Mahoney (1996). The argument that organizational modularity corresponds to product modularity is known as the "mirroring hypothesis." Fine (2000) provides a strong argument as to why firms might want to align their product, process, and supply chain architectures. Sanchez and Mahoney (1996: 64) formulate the mirroring hypothesis as follows: "The [loosely coupled] standardized component interfaces in a modular product architecture provide a form of embedded coordination that greatly reduces the need for overt exercise of managerial authority to achieve coordination of development processes, thereby making possible the concurrent and autonomous development of components by loosely coupled organization structures." The converse argument is that increasing returns to modularity in design can be captured through inter-organizational integration, known as the "complementarity hypothesis." For an analysis in the setting of the Italian air-conditioning industry, see Cabigiosu and Camuffo (2011).

structure, where the lead firm no longer controls the process and has to strive for leadership along with complementary lead firms that supply key components to standardized interfaces. It is within the open-modular phase that platform leadership emerges as a key competitive construct (Gawer, 2009; Gawer & Cusumano, 2002, 2008) and strategies of imitation become significant (Ethiraj, Levinthal, & Roy, 2008). Likewise, the focus on modularity also brings attention back to the cases where integral architectures prove to be superior (Fixson & Park, 2008). The contrast between closed-integral and open-modular architectures is illustrated in Figure 2. In the open-modular case, there is a clear subdivision into sub-assemblies and sub-processes, with minimal interdependence between them. This is the key to their capacity to accelerate diffusion of new technologies.

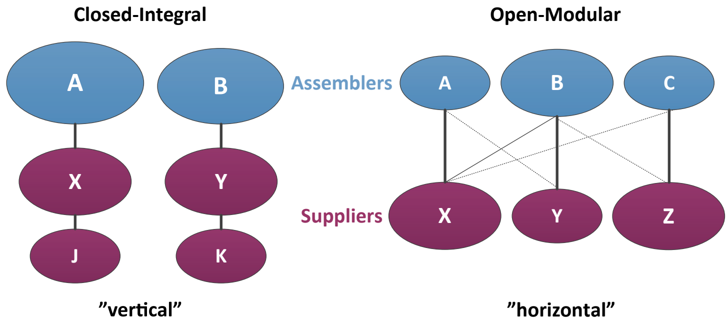

Fig. 2. Closed-integral vs. open-modular industry architectures.
Source: Weinert et al. (2008), Figure 3, adapted from Ge and Fujimoto (2004)

The impact of open-modular industry architecture on competitive dynamics has been the subject of recent analysis. It is generally agreed that the open-modular architecture favors rapid entry by newcomers, particularly from developing countries as the industry moves to low-cost regions. At the same time, the advanced suppliers of core components, such as Intel for the CPU in a PC, or Sanyo Electric for the optical pickup in a DVD player, can make the core component even more integral and hence enhance its competitive advantages. Meanwhile, the scope for upgrading on the part of the developing country firms is considered by some to be limited (e.g., Steinfeld, 2004), but China's ability to leapfrog its way into the global automotive industry, and toward advanced electric vehicles, would seem to cast doubt on this line of argument.[6]

High-technology companies that introduce technologically sophisticated integral products like computers, automobiles, or DVD players have a mixed response to modularization. On the one hand, modularization and standardized interfaces undoubtedly expand the market through cost reduction. But such processes also empower key component suppliers who can wrest control of the platform – as was done by Intel for the hardware and by Microsoft for the software of the IBM PC. And such processes can proceed to the point where latecomer assembly firms from newly industrializing countries can enter the industry, either as modular suppliers (e.g., of motherboards) or as assemblers, as in the case of Taiwanese firms producing the PC at very low cost from standardized modules and creating market opportunities for new firms such as Acer (Mathews & Snow, 1998). When new firms force their way into the industry through utilization of standardized modules, even while incumbent firms continue to compete on the basis of closed-modular or even closed-integral architectures, Fujimoto talks of an intermediate stage called "quasi-open-modular" architecture (Ge & Fujimoto, 2004). We shall discuss such a development in the Chinese two-wheeled electric vehicles case below.

Modularization can enhance opportunities for outsourcing, in the pursuit of increasing returns from specialization. But outsourcing can also work against modularization, as when it is used solely for cost-cutting without regard to modular architecture. Outsourcing simply to

6 Insightful discussion of such issues is currently provided by scholars at the Manufacturing Management Research Center at the University of Tokyo, founded and directed by Takahiro Fujimoto. These scholars see modularity as both a means for latecomer countries to enter new industries, and in response, for integrality to be a viable strategy for Japanese, European, and American leaders, all within a dynamic setting. See Tatsumoto, Ogawa, and Fujimoto (2009) for a discussion of the impact of modularization on the computer industry and the platform strategy pursued by Intel.

capture static cost gains (e.g., by going to lower-wage assembly sites for test and assembly) can actually obstruct the pursuit of modularity (Starr, 2010). When outsourcing is combined with modularity (in a modular production network or MPN, based on manufacturing process outsourcing or MPO), the original equipment manufacturer (OEM) can shift the costs and capital investment risks onto the supply chain, where each supplier is responsible for one of the modules, and each module is allocated to at least one supplier. The suppliers, for their part, benefit from the demand created by the OEM; the supply chain provides such companies with an opportunity to specialize without having to try to cover the range of activities associated with a product.[7] When co-locating firms engage in complementary activities, clusters begin to make their presence felt.

INDUSTRIAL CLUSTERS

Firms in many industries cohere together in various kinds of networks, clusters, and development blocks. There are networks built on networks; indeed, the entire economy can be viewed as interconnected networks of networks (Castells, 2011). Networks grow and become clusters. Such interconnected firm aggregates are well recognized and indeed are becoming the object of increasing attention, due to the outstanding success of such high-tech clusters as Silicon Valley in the USA (Saxenian, 1996) and other science-driven clusters such as Research Triangle Park in North Carolina; the Hsinchu district in Taiwan where all the country's major IT and semiconductor activities are co-located (Mathews & Cho, 2000); and the Jutland region of Denmark (Andersen, 2011). It is widely recognized that the success of regions like Silicon Valley owes much to highly specialized complementarities arising between neighboring firms, something that cannot be accounted for in simple capital and labor terms in a production function.

Clusters and development blocks are the setting in which entrepreneurial and innovative activities can best be understood, as well as the more traditional activities of production of goods and services. This in itself is a powerful departure from earlier traditions that focused on firms acting as individual entities. In the years that have elapsed since Becattini (1990) first pointed out the salience of the Marshallian "industrial district" model to Italy's post-war economic development, and sparked research identifying, counting, and classifying the various kinds of firm concentrations found in Italy and throughout Europe (including famous districts such as the Prato textile district, the Sassuolo ceramic tiles district, and the Carpi knitwear district), the world of scholarship has come to a relatively advanced level of understanding as to what makes industrial clusters work. From the initial enthusiasm linked to the view of districts as being able to generate "flexible specialization" (Piore & Sabel, 1984) through their operations, as alternatives to large integrated firms, there has developed a nuanced understanding of how industrial districts survive and adapt to changing conditions, and how they combine small-firm features with large-firm guidance. The judgment of Harrison (1992) that such districts are not just "new wine in old bottles" remains valid.[8] At the end of

7 Miles and Snow (2007: 459–460) reviewed the field of supply chain management from an organizational perspective, considering the design of multi-firm network organizations, where they conclude that "Supply chain research, which originally focused narrowly on the efficient movement of goods among firms within an industry, now incorporates a substantial amount of organization theory." Further: "The emergence of the multi-firm network organization opened a whole new arena for strategic choice, and many firms became much stronger competitors by linking with specialist providers in an integrated supply chain." On modular production networks, see Sturgeon (2002, 2004).

8 The Florentine scholar Giacomo Becattini is generally regarded as the father figure of Italian industrial district studies, from his 1990 article and earlier work in Italian. His research has been complemented by other notable Italian scholars such as Sebastiano Brusco (1982), who studied the Emilian model, and many others. Harrison (1992) provides an excellent overview that summarizes findings for the 20th century, while Bresnahan, Gambardella, and Saxenian (2001) summarize their studies of emergent "Silicon Valleys" in places such as Ireland, Israel, and Taiwan, identifying five "deep regularities" associated with all such developments: (1) access to highly skilled technical labor; (2) access to managerial labor; (3) institutional forms favoring new firm formation and firm building; (4) connections to markets; and (5) a combination of cooperation with competition. Dahmén (1989) utilized the idea of development blocks, or groups of complementary firms spanning different industries, in his study of the industrialization of Sweden.

the 1990s, Porter (1998, 2000) added his voice to those analyzing and advocating clusters.[9] But the industrial district phenomenon is now seen to encompass not just the advanced world but even more significantly the developing world such as in India, Pakistan, and countries in East Asia.[10] Most notably, the phenomenon can be regarded as the driving factor in China's resurgence as an industrial power, particularly in the setting of the Special Economic Zones (SEZs).

Research on the industrial cluster phenomenon in advanced countries in Europe, North America, and Japan, while developing in varied and important ways such as through investigating the evolution of Marshallian industrial districts and markets-as-networks, has demonstrably failed to diffuse into the disciplines of economics and strategic management, or even organization theory and entrepreneurship studies. Clusters and other supra-firm phenomena such as platforms (Gawer, 2009) and open-modular systems remain on the margins of scholarship despite their real-world significance. It is the rise of industrial clusters in China, and no doubt in India as well, spurred by the creation of Special Economic Zones, which is going to change the situation drastically. My contention is that the success of these emerging industrial giants of the 21st century cannot be understood without reference to the industrial cluster phenomenon that is embedded within them, housed within such institutional settings as SEZs. All the intellectual machinery developed to understand the rise of clusters in the advanced world is now going to have to be applied in order to make sense of this same phenomenon in the developing world, but in a new context defined by globalization and the emergence of global production networks and global value chains (Yeung, 2009). Insights generated through the study of emergent industrial clusters in China and India, and their interaction with global firms and the global value chains that they have been creating, will in turn have repercussions on our understanding as to how such clusters work in the developed world and how they can be created in the developing world – and so the process of mutual scholarly influence will proceed, in a "circular and cumulative causation" pattern that emulates the processes identified for economies more generally.[11]

Industrial clusters are thus widely recognized today to be powerful engines of wealth generation. They may be depicted as microcosms of the economy at large, full of interesting and challenging detail that is passed over by mainstream economics and even by much of strategic management research. How firms enlarge their strategic options through forging connections with one another and in enhancing and deepening the inter-firm knowledge flows that result remains the focus of attention. Firms that form part of a network have access to many more resources than would be available to them individually, and such firms can contract with third parties to accomplish many more activities than would otherwise be under their control, thus expanding the market that is available for their products or services. As the market expands, so the scope for specialization and intermediation grows (exactly as foretold by Adam Smith and earlier by Italian political-economic theorists such as Antonio Serra and Giovanni Botero).[12] This generates a series of positive feedback loops that can be described

9 Michael Porter opened the way to his cluster studies with the identification of a role for "related and supporting industries" in his 1990 framework for analyzing national competitive advantage. But he focused on clusters themselves, particularly at the state level in the U.S. Porter's work has raised public and policy awareness of the significance of industrial clusters.

10 Wei (2009) provides an informative account of the emergence of the Wenzhou model of a footwear cluster in China, while Zhou and Xin (2003) illuminate the ZGC cluster in Beijing. Barbieri, Tommaso, and Huang (2010), Fan and Scott (2003), and Zhu (2009) are examples from a burgeoning literature on Chinese industrial clusters. On the development of Special Economic Zones as an outgrowth of the earlier experiences with Export Processing Zones and Free Trade Zones, first in China and then in India since 2001, see Aggarwal, Hoppe, and Walkinhorst (2009) and Aggarwal (2012). Yusuf, Nabeshima, and Yamashita (2008) provide a useful overview of the extent to which clusters have been "designed" to promote industry development in Asia.

11 The phrase "circular and cumulative causation" was first used by the Swedish development economist Gunnar Myrdal in his 1960 book Asian Drama, and was taken up by the Cambridge economist Nicholas Kaldor as a way of encapsulating real development processes in real economies. See Toner (2000) for a comprehensive discussion.

12 Adam Smith's *Wealth of Nations* (1776) needs no introduction. Italian scholars who anticipated his ideas and elaborated on the role of urban clusters more forcefully must certainly include Antonio Serra (*Breve Trattato delle cause che possono far abbondare li regni d'oro e d'argento dove non sono minere*, 1613: Brief treatise on the causes that can increase wealth in terms of gold and silver where there are no mines) and before him Giovanni Botero (*Delle cause della grandezza delle citta*, 1590: Causes of the greatness of cities). On the significance of their ideas for a long-lost tradition of political economy, but highly relevant to the study of clusters, see Reinert (1999).

as a chain reaction, resulting in the cumulative and circular causation of enhanced production capacities in clusters. Such feedback loops give a spring or bounce to a network that surpasses whatever is available to a firm on its own. The network can reconfigure itself as needed, with inter-firm relations being activated, de-activated, and re-activated as circumstances warrant, leading to a shuffling and reshuffling of the resources embodied in the collective organization. This gives rise to the evolutionary dynamics that generate knowledge spillovers, common resource pools, and interconnections that can then be translated into synergies and systemic returns – better known in economics as increasing returns (Arthur, 1989).

The reshuffling of resources within the cluster may be characterized as an analogue to the reshuffling of the genome of a biological species through Darwinian experimentation and selection, and the shifting activity networks in the cluster that are made possible by this resource reshuffling as the phenotypical expression of these changes in genotype. With due regard to the limitations of biological analogies in the business world, it strikes me that this is a fruitful way to view industrial clusters and to gain insight into the sources of their advantages over the single, isolated firm.[13] Resource dynamics between firms as characterized in this way recalls the principle of biomimesis as the paradigm needed to guide the design of industrial architecture and symbiosis.

CLUSTERING AND MODULARITY

It is not just developing countries like China that utilize modular architectures and clusters to accelerate their market entry and growth. The same strategy can be seen in the case of new market entrants in the developed world – such as Micro Compact Car (MCC), the Mercedes-Swatch joint venture producing the Smart car in Europe. Here it was a case of Swatch bringing its design expertise and the German automaker its production capabilities together in a joint venture created in 1994, and with a new plant built in Hambach, Germany. A cluster of suppliers (known as "system partners") surrounded the MCC plant. Each built one of the modules, such as cockpit, rear axle, or doors, and these modules were then assembled into the final product by the same companies placing their employees in the MCC plant, utilizing a degree of product and process modularization unknown in the automotive industry at the time (Takeishi & Fujimoto, 2001). This was as clear a case of organizational design at the industry level that one is likely to find.

There is no mystery to this process. A comparable process unfolded in Europe in the 1980s with the rise of IKEA as a modular furniture producer and retailer. Prior to IKEA's architectural innovation, the furniture industry was completely closed and integral, with firms supplying their own branded products in vertically integrated fashion. But because the founder of IKEA, Ingvar Kamprad, found himself locked out of the industry by incumbents, he followed a different course and had producers supply modular products exclusively to his own outlet, thus bypassing the closed-integral architecture of the industry. By 2001, IKEA was a global force in furniture retailing and production, with sales of 10.4 billion Euro (US$9.6 billion), a total of 143 company-owned stores in 22 countries, plus another 20 franchised stores and a value constellation of over 2,000 suppliers providing intermediate modular products that IKEA put together in its famous self-assembly kits. How large is the "suction power" of this vast network in the final consumer market? In 2001, there were over 255 million visitors to IKEA stores, and they utilized 110 million catalogs in making purchases. The huge purchasing power assembled by IKEA is what drives the strategizing by the supplier firms, to enroll themselves in the IKEA network. What drives the strategizing by IKEA itself is the platform leadership that enables it to extend its range beyond what any company on its own could expect to accomplish. (I am using the phrase platform leadership in the sense given by Gawer & Cusumano (2002), namely strategizing around the attraction and capture of as many complementary firms as needed to create an industry platform out of a given technology or organizational form.)

This process, where quasi-open modularity is used to accelerate entry into a global business,

13 Nelson and Winter (1982) characterize intra-firm routines as the economic-level analogue of the replicators of a biological system. But the problem is that this notion of routines as replicators does not admit of an easy identification of genotype and phenotype, which is fundamental to the biological conception, whereas to take the argument to the cluster level admits of such an analogue in terms of cluster resources and activities.

driven by the power of a large and growing market to reward specialization, can be found in one industry after another. The process has been studied and documented in the computer, semiconductor, and IT industries; the electronic consumer products industry; the household appliance sector; the furniture sector; and the automotive sector including the two-wheeled vehicles sector, particularly the electric versions of these products. In each case, the process involves a closed-integral product architecture being challenged by a quasi-open modular structure (an architectural innovation, to adopt the terminology of Henderson & Clark (1990)), whereby some firms make the strategic innovation of entering the market as "integrators," based on having a supply of key components from existing industry participants. The new entrants provide an expanded market for the components producers, who are compelled to supply this fresh demand or see new competitors do so. But as they supply the demand, they contribute to the expansion of the market and its further modularization (Jacobides, 2005; Langlois, 2002; Sanchez & Mahoney, 1996). Market expansion creates opportunities for specialized intermediary suppliers to be created, thus driving the process of modularization further. This is exactly the process of market growth and specialization ("division of labor"), reinforcing each other through mutual interaction in a process of circular and cumulative causation, that is expounded by the Smith-Young-Kaldor framework to be discussed below. Modularization is the engine of industry expansion.

The extra twist that I wish to add to this description is how latecomer firms in China and elsewhere in the developing world can make use of these trends toward open modularity (or quasi-open modularity) to facilitate their entry into the industry and accelerate the market expansion through rapid cost reduction. While comparative advantages in terms of low labor costs play an important role, at least in the early stages, it is the competitive advantages created by the approach to industry design, involving clusters and modular networks, that really drive the industries forward. As Wang (2008: 516) puts it, "Architectural innovation is contributing to the Chinese 'catching up' process in terms of industrial development." He argues that local firms opt for a low-cost, low-price strategy as a means of gaining entry, and depending on first-tier and second-tier components suppliers is an optimal way to achieve this. Firms pursue an imitation strategy, engaging in mass production of copied components which come to be standardized at the industry level.

To what extent can we describe this development of a particular kind of industrial organizational architecture (IOA) as the product of design? As an inter-organizational phenomenon, such a quasi-open-modular IOA is clearly beyond the control of any single organization. Yet once it is recognized as a means of promoting latecomer entry into an industry, and accelerating the uptake of green technology, then such industry architecture can be "designed" through national innovation policies that are created to favor such an outcome. And this is precisely what we observe in the case of the Chinese automotive industry and in particular its rapid adoption of electric vehicles, where we see clearly the power of modularity as a principle of design of industrial organizational architecture.

CHINA'S AUTOMOTIVE INDUSTRY AND ELECTRIC VEHICLES

As a latecomer to the automotive industry, China was concerned to promote technology transfer as fast as possible. This it did through a variety of well-known and recognized strategies – foreign direct investment by major Japanese, European, and American producers, with incentives being offered in terms of accessing the Chinese market, in return for commitments to transfer technology and build local supply chains. Modularity undoubtedly played an important role, just as modularity and standardization played an important role in the development of the U.S. automotive industry.[14] The result has been a rapid build-up in China's share of global production, as shown in Figure 3.

14 On the role of modularity and vertical integration vs. "deintegration" in the U.S. automotive industry, see Argyres and Bigelow (2010).

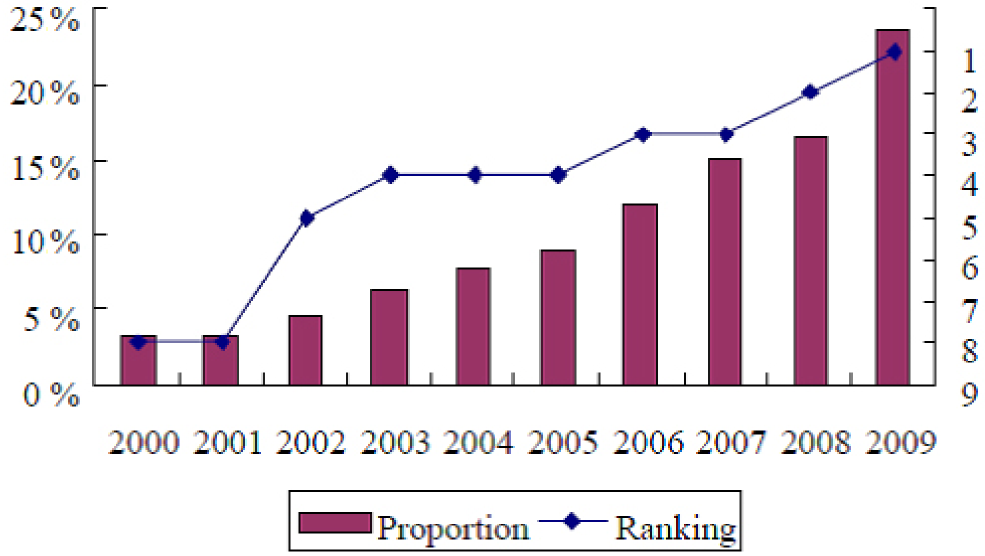

Fig. 3. Global production share and ranking of Chinese automotive firms, 2000–2009.
Source: Wang and Xiao (2011), Figure 1.

The production data on their own tell a fascinating story – that China has invaded the global automotive industry in the space of a decade – and is now set to leapfrog the entire industry to dominate in electric vehicles, both two-wheeled and four-wheeled varieties. Key to this success is modularity. By 2010, China had arrived as the world's largest producer, accounting for 23 percent of global production, up from less than two percent at the beginning of the decade. This extraordinary accomplishment calls for sustained analysis. Many of the Chinese firms had developed beyond mere joint venture partners and were competitive in producing new models of their own. Contributors to this rapid arrival were cost and modularity; in particular, Chinese firms adopted and adapted the platforms that leading foreign producers had created, and which they were induced to transfer to their Chinese partners. There are dozens of these platforms, and Chinese firms have adopted and adapted them all, using them to launch their rapid rise as a global producer. Combined with this adoption of modularity as a driver of international competitive advantage, the Chinese have also promoted clustering in certain key industrial zones such as Shanghai and Guangdong. The two phenomena, modularization and clustering, both reflecting inter-firm industrial dynamics, reinforce each other in synergistic fashion.

While the state-owned firms such as SAIC, FAW, and others have largely remained dependent on their foreign (joint venture) partners (at least in the domain of internal combustion engines), the newer "independent" companies such as Chery, Geely, Great Wall, and BYD have moved ahead rapidly, introducing their own new models with their own brands, by taking advantage of the growing modularity of the industry and the availability of standardized components for the chosen platforms. Some commentators see mere copying in these companies' rise – for example, the Chery QQ being based on the Daewoo Matiz, or the Great Wall Peri based on the Fiat Panda, or the Chinese "jeeps" (high-mobility multi-purpose wheeled vehicles HMMPWVs for military use) built by Dongfeng Motor being based on components from AM General (formerly American Motors). But the key to their success is their capacity to drive the process of modularization from quasi-open modular to fully open modular, in a process that resembles what Tatsumoto et al. (2009) call the modular "separation effect."

Geely's transformation is emblematic of this gap effect. Founded in 1986 as a manufacturer of refrigerators, Geely moved into production of motorcycles in 1994 and then into vehicle production in 1998. At that time its vehicles had to be classified as "vans" to avoid certification issues, but Geely received state approval to manufacture automobiles in 2001. At that time there were upward of 120 producers in China (official and unofficial), and government policy favored the big state-owned corporations like FAW and SAIC that were locked into joint ventures with foreign partners. Geely began by mixing and matching existing components.

Its first independent model was the *Haoqing*, based on the Charade, produced in China by FAW, the result of technology transfer from Toyota's affiliate, Daihatsu. No fewer than 70 percent of the components utilized were interchangeable with the Charade – 60 percent supplied directly from FAW and a further 10 percent based on copied standardized units.[15] The next model, the *Maple*, introduced in 2002, was based on interchangeable parts from two foreign models already produced in China, the Citroen ZX and again the Charade. This represented an "architectural leap" for Geely, which in itself helped to drive the industry to a quasi-open modular architecture. Subsequently, Geely has produced engines that can fit within several extant platforms, while Geely's cars can accommodate engines bought from other manufacturers, all based on standardized interfaces. Complementing the quasi-open modularity of its product, Geely has designed a modular supply chain, outsourcing the majority of components to suppliers with whom it works closely and has long-term relations. In this way, Geely was able to build on its links with suppliers for its motorcycle business, incorporating suppliers who also supply to the big foreign companies such as VW, Toyota, Nissan, and GM. As Wang and Kimble (2010b: 18) state: "Having Geely as the customer helps those suppliers realize economies of scale because of the higher production volume."

The same argument can be made with regard to other successful Chinese independent producers such as Chery Automobile. Chery is a small state-owned company based in Wuhu, 200 km west of Shanghai, founded in 1997, that built its first prototype self-badged car in 1999, and has since expanded rapidly, becoming China's largest vehicle exporter by 2007. The first prototype was based on a chassis licensed from Volkswagen. The key to its rapid insertion in the international industry is its preparedness to tap into the international supply base created by an increasingly modular automotive industry. This style of development has been aptly called "compressed development" and modularity, particularly its open version, would seem to drive it fastest.[16]

The proportion of Chinese cars featuring quasi-open modular product architecture has now reached probably 30 percent of total production in China – a proportion that is growing and which is having a much greater influence than the closed-integral architecture traditionally employed by the automotive majors. Component suppliers themselves are helping to drive this process, merely by responding to the growing demand. For example, Mitsubishi now sells engines to at least 21 carmakers in China, while the engine management system module produced by Delphi (formerly linked to GM) can be utilized with any Mitsubishi engine. In the same way, makers of DRAMs in the semiconductor industry were driven to make their modular parts standardized and consistent with Intel's CPU chipsets, thereby accelerating the rate of market expansion (Tatsumoto et al., 2009).

Electric Vehicles

Most recently, China has formulated strategies designed to leapfrog the world's automotive giants and become the leading producer of electric vehicles (EVs). Originally, China focused its New Energy Vehicles on alternative fuels, particularly diesel, and then on electric hybrids that run on both their gasoline engine and their electric motors, powered by their on-board batteries and recharging systems (such as regenerative braking). But with the 12th Five-Year Plan (covering the years 2011–2015), China is clearly focused on pure electric vehicles (i.e., on battery-powered EVs) and plug-in hybrid electric vehicles (PHEVs), which are seen as the next wave for zero-emission vehicles. They are also much simpler in design and construction, since there is either no internal combustion engine or it is operated in ancillary mode only. Chinese automotive firms are leaping (indeed leapfrogging) to adopt these new forms of

15 This section is based on Wang (2008) and Wang and Kimble (2010a, 2010b, 2011).

16 On compressed development, and for further discussion of the Chery case, see Whittaker, Zhu, Sturgeon, Tsai, and Okita (2010). Sturgeon and Van Biesenbroeck (2010: 13) describe Chery as follows: "Chery Automobile, a small, state-controlled company based in Wuhu, China, has been able to develop and market a line of Chery brand vehicles within a remarkably short time by tapping the expertise of first-tier global suppliers with operations both in China and in the West. Chery obtains a full range of inputs from the global supply base, from parts to production equipment to design and system integration expertise." They then qualify the point, stating that "Since learning is relatively shallow, the sustainability of Chery's approach will need to be proven over the long term." Their analysis does not move to the point where Chery and other Chinese vehicle producers have leapfrogged to the new paradigm of the electric vehicle.

EVs and PHEVs, building their own models from platforms adapted from those already developed in the automotive industry generally. The progress registered so far is impressive.

The FAW group (formerly First Auto Works), now one of China's largest automotive producers (more than 1 million units annually) has already built two new EV production facilities, in Changchun and Dalian city. The newly launched vehicles, including the Besturn 50 plug-in hybrid electric vehicle (PHEV) and Besturn 70 electric vehicle (EV), were independently developed by FAW Group. The company asserts that it will invest a total of 10 billion yuan to develop eight kinds of new energy vehicle product platforms, leading to 13 new energy passenger cars and three commercial vehicles during the Twelfth Five-Year Plan period.[17] FAW now appears to be thoroughly emancipated from its earlier dependence on Mazda.

The private producer Geely has introduced its electric vehicle, the Nanoq (meaning "polar bear" in Greenland), to be supplied in Europe through a joint venture with the Danish company Lynx. The joint venture will utilize batteries supplied by Lynx GT (described below). Other producers such as Changan, Chery, and SAIC are introducing their own models, based on modular platforms that speed up the process of development. They are rapidly building modular supply chains in the form of clusters to support their EV production activities. The key module in the case of EVs is the battery itself, and here there has been super-charged activity by Chinese firms linking themselves to European and American firms. The support provided by the government-owned banks channeling investment into the sector means that the electric vehicle industry is growing fast, with modular organizational architecture driving the development.

The exception to this trend is BYD, one of the spectacular success stories of the EV sector. A major lithium-ion battery producer for the cellular phone and consumer goods sector, founded in 1995, it elected to enter the automotive industry by acquiring an existing automotive producer. By 2008, BYD was able to produce its first hybrid electric (or dual mode) vehicle, the F3DM, which started shipping in early 2010, then a four-wheel drive S6DM (with dual motors controlling the front wheels and a 75kW electric motor the back wheels). In 2010, it launched a pure EV, the e6, a five-seater sedan with a 75kW electric motor and the BYD proprietary Fe lithium-ion battery. BYD is working with taxi operators and municipal governments in China to grow the market for its EVs, and in 2010 formed a joint venture with the German firm Daimler to produce EVs for China under the joint venture's own brand. BYD follows a quite different strategy from other Chinese automotive producers in seeking to control its entire value chain through vertical integration. BYD even produces its own charging stations, giving it a unique systemic perspective on the entire EV industry.

China is also paying close attention to the building of the infrastructure needed to support electric vehicles. It shows marked originality in the fact that the Chinese overseas oil company CNOOC (China National Offshore Oil Corporation) is investing in the acquisition of domestic petrol stations with a view to turning them into EV charging centers – in competition with the domestic oil firms Sinopec and CNPC. Likewise, the electric power grid companies, State Grid (SGCC) and China Southern Power Grid (CSG), are looking to build charging networks and integrating them into their planned smart grid developments. Again, JVs are being utilized to accelerate technology diffusion. In 2011, Better Place, the American-Israeli company producing EV's battery charging and switching stations, announced an agreement with CSG envisaging the opening of battery switch stations by the end of the year.

Under China's 12th Five-Year Plan, electric vehicles (called "New Energy Vehicles") are designated as one of the seven strategic industries to be promoted through a range of tax breaks, subsidies, and technology promotion. Under the Plan, an interim target of 1 million vehicles is set for 2015, with a longer-term goal of reaching 100 million vehicles (cars and commercial vehicles) by 2020. One can agree with Wang and Kimble (2011) that this is indeed a national leapfrogging strategy.[18]

17 See "FAW's New Energy Vehicle Launches August 23, 2011" at: http://www.faw.com/faw_online/news/dzjy_jybj/jyzb/20110829133200038.htm

18 Sources on China's EV initiatives include Weinert et al. (2008) and Wang and Kimble (2010a,b; 2011) as well as Wang and Xiao (2011).

Two-Wheeled Electric Vehicles

The most spectacular effect of modularity is seen in the case of two-wheeled EVs, or electric "motor" bicycles (E2Ws). Here the impact of modularity is linked to market promotion via local municipalities banning the use of internal combustion engine motorcycles in cities on account of the noise and fumes they emit. The result has been a marked rise in production and sales of Chinese electric two-wheeled vehicles, which by 2007 had reached a level approaching 15 million units produced per year and far exceeding production of all passenger vehicles (see Figure 4). By 2010, production was estimated to reach 30 million units. Here is an industry that cries out for close study from an organizational perspective.

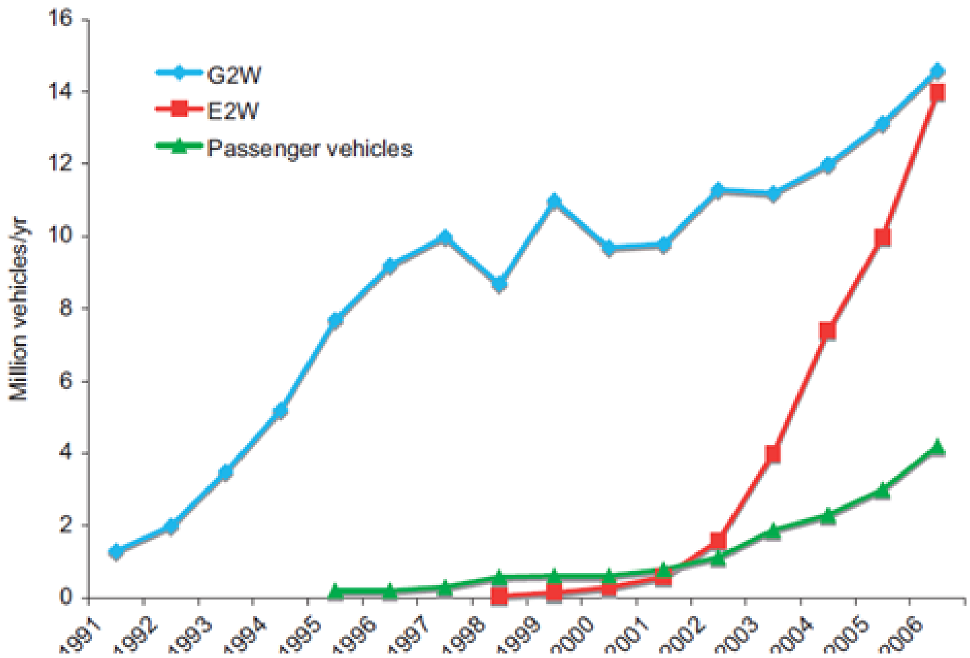

Fig. 4. Growth of electric two-wheeled vehicles in China, 2000–2007.
Source: Weinert et al. (2008).

Just like the bicycle industry before it, the e-bicycle industry is highly modular, with key components like batteries supplied by just a few critical suppliers. In terms of the Fujimoto classification, the E2W industry in China is quasi-open modular (Ge & Fujimoto, 2004). The e-bicycles are assembled from just a few components, which because of high levels of demand have become rapidly standardized. Each final producer can work with multiple component suppliers, while each module producer can sell to multiple assemblers. Batteries come from firms such as Protanium, the lithium-ion battery producer. And again, just as the bicycle industry has demonstrated some cyclicality in its "reverse integration," moving from extreme modularity to closer integration (via the Japanese firm Shimano), so the E2W industry in China is also showing the two trends at work. While the industry becomes more fragmented and modularized, at the same time a leading company, Xinri, is building a distinctive competitive advantage through vertical integration, now covering every step in the E2W value chain except batteries, and emerging as China's largest supplier. Xinri's base is in Wuxi, where a vast modular supply cluster is developing to power the E2W sector.[19]

ROLE OF SUPRA-FIRM STRUCTURES IN INDUSTRY GROWTH AND STABILITY: AN ANALYTIC FRAMEWORK

Let me draw the threads of this discussion together by focusing on the key drivers of the expansion of clusters and modular systems within the wider economy. There is a rich "heterodox" economics tradition that organization and strategy theorists have yet to tap into in explaining the growth and diffusion of the modular and open inter-organizational

19 See the company's website at: http://english.xinri.com/

architectures discussed so far. Let me mention just three: (1) Allyn Young, in his 1928 Address to the British Association for the Advancement of Science, on the theme of specialization (division of labor) and increasing returns; (2) Alfred Marshall, in his notion of firms within clusters generating increasing returns through external economies; and (3) David Ricardo, in his idea that two countries can specialize and improve their economic position by the resultant "gains from trade."

It was Young (1928) who boldly posed the issue of increasing returns as the central question to be addressed in economic analysis of the modern industrial system. In place of seeing the genesis of increasing returns as a marginal issue, to be dealt with alongside externalities as something quaint and uncommon, Young grasped that increasing returns are central to the way that mass-production industries go about building the market for their products. On the strength of the expanded market, they are able to invest in specialized capital equipment, and as the market further expands they are able to make use of specialized value chains of intermediate suppliers, sometimes aggregated together in industrial clusters. Young insisted that it is not factor questions and supply-side issues that need to be addressed in accounting for increasing returns but growth of markets (i.e., growth in demand). This demand-side emphasis is a singular characteristic of Young's framework, an emphasis that mainstream economics has ignored. Young insisted that firms in modern mass-production industries first address the market and take active steps to build the market prior to making definitive investments in production. The other feature of such firms' investment behavior, which again did not escape the notice of Young, is their preparedness to sink large sums into investment in large-scale production systems that would be completely unwarranted by the current state of demand. Such investments are made with an eye on growing the market through cost reduction as fast as possible. Cost reductions are based on prior investments in specialized capital equipment provided by specialist suppliers whose existence is made possible by the breadth of the market as well as in internal efficiencies that are under the firm's direct control. It is the efficiency gains on the part of external suppliers that generate Marshall's "external economies."[20]

To translate into the language of modularity, Young's account makes sense of why modular production systems may outperform closed, integral systems (up to a point) because the appearance of specialized stable intermediaries enables final producers to reap productivity gains (increasing returns) while the productivity improvements of the stable intermediaries (modular suppliers in a value chain) drive expansion of the market. As the market expands, it creates opportunities for even more specialized intermediaries to appear, and as they improve productivity, they enable the market to further expand, and so on – in a process of circular and cumulative causation that is best described as a chain reaction. This is why modular industrial systems will outperform any other industrial architecture, if they are allowed to express themselves and if they are initiated by deliberate, entrepreneurial acts of industrial design and by supportive government policies.

Indeed, there is an opportunity here to adapt Ricardo's 19$^{\text{th}}$-century doctrine of comparative advantage, as formulated between countries.[21] Ricardo's doctrine is that two countries, A and B, will improve their position (income) by trading goods in which they specialize (i.e., where they possess a comparative advantage). Now transpose this argument to the case of two firms in a value chain. Firms A and B will improve their joint productivity by sourcing from each other's specialization. Here surely lies the origin of economists' general category of "increasing returns." The gains from trade are real and can be observed at the firm level.

As firms increase their level of specialization, and play the role of specialized suppliers of intermediate products, they create multiple opportunities to enhance their joint productivity through multiple inter-firm interactions. This is precisely what Marshall (1890) was pointing to in his rather vague formulation of "external economies" – that is, the savings reaped by firms from their interaction with neighboring firms as these neighboring firms improve their productivity through specialization. Thus, the most obvious case of Ricardo's gains from trade would be Marshall's external economies reaped by firms in an industrial cluster through their interactions. In both cases, modularity and clustering, we observe firms benefiting from

20 See Marshall's original exposition in his *Principles of Economics* (1890).

21 See Ricardo's original exposition in his *Principles of Political Economy and Taxation* (1821).

systemic gains arising from inter-firm activities that would not be available to a firm acting on its own. And, as asserted by Young, it is market expansion that triggers the appearance of these specialized suppliers of intermediates, and their appearance then drives further expansion which triggers further specialization, and so on. What this process generates is economic expansion and increasing returns.

Some strategic management and organizational scholars have linked gains from trade with the inter-organizational architecture of value chains.[22] There is clearly scope here to derive gains from trade at the inter-firm level as a strategic consequence of different forms of inter-organizational industry architecture, and to link such phenomena to the wider picture created by Ricardo's comparative advantage, Marshall's external economies, and Young's increasing returns and division of labor. In this way, we can fashion the theoretical framework that explains why firms organize themselves into value chains and allow modularization (through specialized componentry) to appear, and what drives the entire process.[23]

CONCLUSION

In this article, I have sought to demonstrate the power and salience of supra-firm structures in the design of industries, in particular the role played by modularity (and degrees of modular openness) in accounting for the success and failure of industries so created. The emphasis has been on the role played by modularity in accelerating the uptake of new technologies and enhancing opportunities for firms to break into established markets or create new ones. The utility of such a framework is demonstrated by examining the case of China's entry into industries such as automobiles, and into newly emerging sectors such as two-wheeled and four-wheeled electric vehicles. Here conventional explanations of China's successes, which focus on costs and government coordination and macroeconomic manipulation, can account for only part of the success. The less obvious but arguably more important features have to do with the power of modularity and its influence on the speed of diffusion of new product forms and patterns of industrial organization. Clusters, networks, and modularity are the all-important organizational features essential to understanding these recent market upheavals.

The capacity of modular organizational systems to accelerate diffusion of new technologies and facilitate market entry, and the "designability" of such industrial structures, have an immediate and important bearing on the greening of the industrial model being adopted by China and India. The issue of global warming is addressed by a variety of scientific disciplines working within a conventional policy framework. Mitigation of carbon emissions is ritually called for in public statements, but how this is to be achieved without blocking the industrialization efforts of new industrial giants such as China, India, and Brazil has not been satisfactorily resolved. Promotion of renewable energies, through market expansion programs and cost reduction strategies, are widely supported but meet stiff political and economic resistance from fossil fuel lobbies who see their interests threatened.

The promotion of specific kinds of supra-firm organizational architectures, such as clusters and quasi-open-modular structures, provide the missing ingredient, both in development strategies as well as in market-entry strategies by advanced firms. It is the close attention paid to such matters in China that helps explain the country's remarkable rise as a manufacturing power over the course of the past decade and as a source of much of the green technology that is diffusing around the world.

The introduction of new inter-organizational industrial architectures along the lines discussed here – designing open-modular architectures wherever feasible, promoting clustering and inter-firm linkages to turn wastes into inputs – promises to provide a new avenue for bringing industrial capitalism into alignment with its ecological setting and providing industry with a biomimetic paradigm to guide its growth and sustainability. Such

22 Jacobides and Winter (2005), for example, contrast two industries – the U.S. mortgage banking industry (prior to the bursting of the subprime lending bubble) and the Swiss watch industry – in terms of their fluctuating degrees of modularity. In the process they discuss how firms benefit from sourcing components to specialized suppliers, capturing gains from trade. Jacobides and Hitt (2005) provide a similar reference to gains from trade where this is interpreted in terms of firms with different capabilities along a value chain either complementing each other or integrating their operations.

23 Aspects of this issue have been tackled in the theoretical economics literature, notably by Yang and Borland (1991), where they investigate increasing returns to specialization in a dynamic general equilibrium model.

organizational design approaches to climate change mitigation do not appear to have registered as yet with global policy bodies. It is surely time to widen the frame of thinking and to give organizational and strategy scholars an opportunity to make a real contribution to redesigning the capitalist economic processes that give the global industrial system such vibrancy and wealth-generating potential. The susceptibility of particular industries to modular redesign, and their capacity to support rapid diffusion of new technologies, especially low-carbon, clean technologies, is a critical public issue where scholarship and policy development will have to move together in unison.

REFERENCES

Aggarwal A. 2012. *Social and Economic Impact of SEZs in India*. Oxford University Press, Oxford, UK.

Aggarwal A, Hoppe M, Walkenhorst P. 2009. Special Economic Zones and economic diversification: Some evidence from South Asia. In R. Newfarmer, W. Shaw, and P. Walkenhorst (Eds.), B*reaking into New Markets: Emerging Lessons for Export Diversification*. The World Bank, Washington, DC.

Andersen PH. 2011. Cluster-based global firms use of local capabilities. *Management Research Review* 34(10): 1087–1106.

Argyres N, Bigelow L. 2010. Innovation, modularity, and vertical deintegration: evidence from the early U.S. auto industry. *Organization Science* 21(4): 842–853.

Arthur B. 1989. Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal* 99(394): 116–131.

Baldwin CY, Clark KB. 2000. *Design Rules, Volume 1: The Power of Modularity*. MIT Press, Cambridge, MA.

Baldwin CJ, Woodard CJ. 2007. Competition in modular clusters. Working paper, School of Information Systems, Singapore Management University.

Baldwin CJ, Woodard CJ. 2009. The architecture of platforms: A unified view. In A. Gawer (Ed.), *Platforms, Markets and Innovation*: 19–44. Edward Elgar, Cheltenham, UK.

Barbieri E, Di Tommaso M, Huang M. 2010. Industrial development policy and innovation in southern China: Government targets and firms' behaviour. *European Planning Studies* 18(1): 83–105.

Becattini G. 1990. The Marshallian industrial district as a socioeconomic notion. In F. Pyke, G. Becattini, and W. Sengenberger (Eds.), *Industrial Districts and Inter-Firm Cooperation in Italy*. Geneva: International Institute for Labour Studies.

Benyus JM. 1997. *Biomimicry: Innovation Inspired by Nature*. William Morrow, New York.

Benyus JM, Baumeister D. 2004. *Principles of Life*. Biomimicry Guild.

Bresnahan TF, Gambardella A, Saxenian, A. 2001. 'Old economy' inputs for 'new economy' outcomes: cluster formation in the new Silicon Valleys. *Industrial and Corporate Change* 10(4): 835–60.

Brusco S. 1982. The Emilian model: Productive decentralization and social integration. *Cambridge Journal of Economics* 6(2): 167–184.

Brusoni S, Prencipe A. 2001. Unpacking the black box of modularity: technologies, products and organizations. *Industrial and Corporate Change* 10: 179–205.

Brusoni S, Prencipe A, Pavitt K. 2001. Knowledge specialization, organizational coupling, and the boundaries of the firm: why do firms know more than they make? *Administrative Science Quarterly* 46(4): 597–621.

Cabigiosu A, Camuffo A. 2011. Beyond the 'mirroring hypothesis': product modularity and interorganizational relations in the air conditioning industry. *Organization Science* (in press).

Campagnolo D, Camuffo A. 2010. The concept of modularity in management studies: a literature review. *International Journal of Management Reviews* 12(3): 259–283.

Castells M. 2011. *The Rise of the Network Society*. Wiley, New York.

Chesbrough H. 2005. Towards a dynamics of modularity: a cyclical model of technical advance. In A. Prencipe, A. Davies, and M. Hobday (Eds), *The Business of Systems Integration*. Oxford University Press, New York.

Chesbrough H, Kusunoki K. 2001. The modularity trap: technology, innovation, technology phase shifts, and the resulting limits of virtual organizations. In I. Nonaka and D. Teece (Eds), *Managing Industrial Knowledge: Creation, Transfer and Utilization*: 202–230. SAGE, London, UK.

Chesbrough H, Prencipe A. 2008. Networks of innovation and modularity: a dynamic perspective. *International Journal of Technology Management* 42(4): 414–425.

Dahmén E. 1989. Development blocks in industrial economics. In B. Carlsson (Ed.), *Industrial Dynamics*. Kluwer Academic Publishers, Boston/Dordrecht.

Ethiraj SK, Levinthal D, Roy RR. 2008. The dual role of modularity: innovation and imitation. *Management Science* 54: 939–955.

Fan CC, Scott AJ. 2003. Industrial agglomeration and development: a survey of spatial economic issues in East Asia and a statistical analysis of Chinese regions. *Economic Geography* 79(3): 295–319.

Fine CH. 2000. Clockspeed-based strategies for supply chain design. *Production and Operations Management* 9(3): 213–221.

Fine CH, Golany B, Naseraldin H. 2005. Modeling tradeoffs in three-dimensional concurrent engineering: a goal programming approach. *Journal of Operations Management* 23: 389–403.

Fixson SK, Park JK. 2008. The power of integrality: linkages between product architecture, innovation, and industry structure. *Research Policy* 37: 1296–1316.

Fligstein N. 1985. The spread of the multidivisional form among large firms, 1919–1979. *American Sociological Review* 50(3): 377–381.

Franko L. 1976. *The European Multinationals*. Greylock Press, Greenwich, CT.

Fujimoto T. 1999. *The Evolution of a Manufacturing System at Toyota*. Oxford University Press, New York.

Fujimoto T. 2006. Architecture-based comparative advantage in Japan and Asia. MMRC Discussion paper 94. Manufacturing Management Research Center, University of Tokyo.

Garud R, Kumaraswamy A. 1995. Technological and organizational designs for realizing economies of substitution. *Strategic Management Journal* 16: 93-109.

Gawer A. (Ed.) 2009. *Platforms, Markets and Innovation*. Edward Elgar, Cheltenham, UK.

Gawer A, Cusumano MA. 2002. *Platform Leadership: How Intel, Microsoft and Cisco Drive Industry Innovation*. Harvard Business School Press, Boston, MA.

Gawer A, Cusumano MA. 2008. How companies become platform leaders. *MIT Sloan Management Review* 49: 28–35.

Ge DS, Fujimoto T. 2004. Quasi-open product architecture and technological lock-in: an exploratory study of the Chinese motorcycle industry. *Annals of Business Administrative Science* 3(2): 15–24.

Gore A. 2007. *An Inconvenient Truth: The Crisis of Global Warming*. Bloomsbury Publishing, London, UK.

Harrison, B. 1992. Industrial Districts: old wine in new bottles? *Regional Studies* 26(5): 469-483.

Henderson RM, Clark KB. 1990. Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35: 9–30.

Jacobides MG. 2005. Industry change through vertical disintegration: how and why markets emerged in mortgage banking. *Academy of Management Journal* 48(3): 465–498.

Jacobides MG, Hitt LM. 2005. Losing sight of the forest for the trees? Productive capabilities and gains from trade as drivers of vertical scope. *Strategic Management Journal* 26: 1209–1227.

Jacobides MG, Winter SG. 2005. The co-evolution of capabilities and transaction costs: explaining the institutional structure of production. *Strategic Management Journal* 26: 395–413.

Langlois RN. 2002. Modularity in technology and organization. J*ournal of Economic Behavior and Organization* 49: 19–37.

Langlois RN. 2003. The vanishing hand: the changing dynamics of industrial capitalism. *Industrial and Corporate Change* 12(2): 351–385.

Langlois RN, Robertson PL. 1992. Networks and innovation in a modular system: lessons from the microcomputer and stereo component industries. *Research Policy* 21: 297–313.

Marshall A. 1890. *Principles of Economics*. Macmillan, London, UK.

Mathews JA. 1996. Holonic organizational architectures. *Human Systems Management* 15(1): 27–54.

Mathews JA. 2002a. The origins and dynamics of Taiwan's R&D consortia. *Research Policy* 31(4): 633–651.

Mathews JA. 2002b. *Dragon Multinational: A New Model of Global Growth*. Oxford University Press, New York.

Mathews JA. 2003. Competitive dynamics and economic learning: an extended resource-based view. *Industrial and Corporate Change* 12(1): 115–145.

Mathews JA. 2006a. *Strategizing, Disequilibrium and Profit*. Stanford University Press, Stanford, CA.

Mathews JA. 2006b. Ricardian rents or Knightian profits: more on Austrian insights on strategic organization. *Strategic Organization* 4(1): 97–108.

Mathews JA. 2010. Lachmannian insights into strategic entrepreneurship: resources, activities and routines in a disequilibrium world. *Organization Studies* 31(2): 219–244.

Mathews JA, Cho DS. 2000. *Tiger Technology: The Creation of a Semiconductor Industry in East Asia*. Cambridge University Press, Cambridge, UK.

Mathews JA, Snow CC. 1998. A conversation with Stan Shih on global strategy and management. *Organizational Dynamics* (Summer): 65–74.

Mathews JA, Tan H. 2011. Progress towards a circular economy in China: drivers (and inhibitors) of eco-industrial initiative. *Journal of Industrial Ecology* 15(3): 435–457.

Miles RE, Snow CC. 2007. Organization theory and supply chain management: an evolving research perspective. *Journal of Operations Management* 25(2): 459–463.

Miles RE, Snow CC, Mathews JA, Miles G, Coleman HJ. 1997. Organizing in the knowledge age: anticipating the cellular form. *Academy of Management Executive* 11(4): 7–20.

Nelson RR, Winter SG. 1982. *An Evolutionary Theory of Economic Change*. The Belknap Press of Harvard University Press, Cambridge, MA.

Ogawa K, Shintaku J, Yoshimoto T. 2005. Architecture-based advantage of firms and nations: new global alliance between Japan and catch-up countries. MMRC Discussion Paper 48. Manufacturing Management Research Center, University of Tokyo.

Piore M, Sabel C. 1984. *The Second Industrial Divide: Possibilities for Prosperity*. Basic Books, New York.

Polanyi K. 1944 (1957, 2001). *The Great Transformation: The Political and Economic Origins of Our Time*. Beacon Press, Boston, MA.

Pomeranz K. 2000. T*he Great Divergence: China, Europe and the Making of the Modern World Economy*. Princeton University Press, Princeton, NJ.

Porter ME. 1998. Clusters and the new economics of competition. *Harvard Business Review* (Nov/Dec): 77–90.

Porter ME. 2000. Location, competition, and economic development: local clusters in a global economy. *Economic Development Quarterly* 14(1): 15–34.

Reap J, Baumeister D, Bras B. 2005. Holism, biomimicry and sustainable engineering, *Proc IMECE2005*, 2005 ASME International Mechanical Engineering Conference, 8134, pp. 1–9.

Reinert ES. 1999. The role of the state in economic growth. *Journal of Economic Studies* 26(4/5): 268–326.

Ricardo D. 1821. *On the Principles of Political Economy and Taxation*. John Murray, London, UK.

Sanchez R, Mahoney JT. 1996. Modularity, flexibility, and knowledge management in product and organization design. *Strategic Management Journal* 17(Winter Special Issue): 63–76.

Saxenian A. 1996. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.

Schilling MA. 2000. Toward a general modular systems theory and its application to interfirm product modularity. *Academy of Management Review* 25: 312–334.

Simon HA. 1962. The architecture of complexity. *Proceedings of the American Philosophical*

Society 106(6): 467–482.
Spence M. 2011. T*he Next Convergence: The Future of Economic Growth in a Multispeed World*. Farrar, Straus and Giroux, New York.
Starr MK. 2010. Modular production – a 45-year old concept. *International Journal of Operations & Production Management* 30(1): 7–19.
Steinfeld ES. 2004. China's shallow integration: networked production and new challenges for late industrialization. *World Development* 32(11): 1971–1987.
Stern, N. 2007. *The Stern Review of the Economics of Climate Change*. London: H.M. Treasury. Available at: http://www.hm-treasury.gov.uk/sternreview_index.htm
Sturgeon TJ. 2002. Modular production networks: a new American model of industrial organization. *Industrial and Corporate Change* 11(3): 451–496.
Sturgeon TJ, Van Biesenbroeck J. 2010. Effects of the crisis on the automotive industry in developing countries: a global value chain perspective. Policy Research Working Paper 5330. The World Bank, Washington, DC.
Takeishi A, Fujimoto T. 2001. Modularisation in the auto industry: interlinked multiple hierarchies of product, production and supplier systems. *International Journal of Automotive Technology and Management* 1(4): 379–396.
Tatsumoto H, Ogawa K, Fujimoto T. 2009. The effect of technological platforms on the international division of labor: a case study of Intel's platform business in the PC industry. In A. Gawer (Ed.), *Platforms, Markets and Innovation*: 345–370. Edward Elgar, Cheltenham, UK.
Toner P. 2000. *Main Currents in Cumulative Causation: The Dynamics of Growth and Development*. Palgrave Macmillan, London, UK.
Voordijk H, Meijboom B, De Haan J. 2006. Modularity in supply chains: A multiple case study in the construction industry. *International Journal of Operations and Production Management* 26(6): 600–618.
Wang H. 2008. Innovation in product architecture: a study of the Chinese automobile industry. *Asia Pacific Journal of Management* 25: 509–535.
Wang H, Kimble C. 2010a. Betting on Chinese electric cars? Analysing BYD's capacity for innovation. *International Journal of Automotive Technology and Management* 10(1): 77–92.
Wang H, Kimble C. 2010b. Low-cost strategy through product architecture: Lessons from China. *Journal of Business Strategy* 31(3): 12–20.
Wang H, Kimble C. 2011. Leapfrogging to electric vehicles: Patterns and scenarios for China's automobile industry. *International Journal of Automotive Technology and Management* 11(4): 312–325.
Wang HL, Xiao J. 2011. Delivering discontinuous innovation through modularity: The case of Chinese electric vehicle industry. *Proceedings ICMET conference*.
Wei YHD. 2009. China's show manufacturing and the Wenzhou model: Perspectives on the world's leading producer and exporter of footwear. *Eurasian Geography and Economics* 50(6): 720–739.
Weinert J, Ogden J, Sperling D, Burke A. 2008. The future of electric two-wheelers and electric vehicles in China. *Energy Policy* 36: 2544–2555.
Whittaker DH, Zhu T, Sturgeon T, Tsai MH, Okita T. 2010. Compressed development. *Studies in Comparative International Development* 45: 439–467.
Yang XK, Borland J. 1991. A microeconomic mechanism for economic growth. *Journal of Political Economy* 99: 460–482.
Yeung HWC. 2009. Regional development and the competitive dynamics of global production networks: an East Asian perspective. *Regional Studies* 43(3): 325–351.
Young AA. 1928. Increasing returns and economic progress. *The Economic Journal* 38: 527–542.
Yusuf S, Nabeshima K, Yamashita S. (Eds.) 2008. *Growing Industrial Clusters in Asia: Serendipity and Science*. The World Bank, Washington, DC.
Zhou Y, Xin T. 2003. An innovative region in China: interaction between multinational corporations and local firms in a high-tech cluster in Beijing. *Economic Geography* 79(2): 129–152.

Zhu YM. 2009. *Analysis of Industrial Clusters in China*. CRC Press, Baton Rouge, LA.

JOHN A. MATHEWS

Chair of Strategic Management
Macquarie Graduate School of Management, Macquarie University, Sydney
E-mail: john.mathews@mgsm.edu.au

DESIGNING ORGANIZATIONS FOR EXPLORATION AND EXPLOITATION

TIMOTHY N. CARROLL

Abstract: All organizations face the core challenge of deciding on investments in two very different types of activities: exploration and exploitation. Exploration activities are future-oriented, such as developing new capabilities, experimenting with new technologies, and pursuing new customers and markets. Exploitation activities, in contrast, focus on the refinement of existing competencies, processes, and products. Because an organization's design should reflect its goals, it is difficult to accommodate exploration and exploitation activities within a single organization. This article discusses four major approaches used to tackle this problem, and notes the strengths and limitations of each approach.

Keywords: Exploration-exploitation; organizational ambidexterity; dynamic capabilities; organization design

One of the most influential scholarly works on organizations over the last twenty years has been James March's "Exploration and Exploitation in Organizational Learning" (March, 1991). Part of the article's appeal is that it addresses a core organizational challenge – deciding between investments in two different types of activities. Pursuing exploitation activities implies a focus on the "refinement and extension of existing competencies, technologies and paradigms," while an exploration focus indicates "experimentation with new alternatives" (March, 1991: 85). Exploitation is necessary for improving current operations, and returns on investments in exploitation are likely to be near term and positive. Exploration is more likely to yield the next breakthrough idea, product, or market, but returns on exploration are less certain and more distant in time. The exploration-exploitation construct can be applied to organizational choices related to alliances, new product development, which markets and customers to target, and employee and organizational development.

Although exploration and exploitation are both important to organizational performance, most organizations would like to be able to pursue each type of activity at the same time. Doing both simultaneously, however, can be difficult. For one thing, each approach can become self-reinforcing. Exploration, for example, is by its very nature variable and prone to failure. When inevitable failures provoke a search for other new approaches, the organization may fall prey to the "failure trap" – always looking for the next great thing. Such organizations pay the costs of experimentation without gaining the benefits. Conversely, since an exploitation approach is more likely to yield early successes, these can reinforce the pursuit of similar efforts, creating a "success trap." While this promotes stability, it also keeps the organization from finding new opportunities. A second challenge is that competitive pressures may push the organization to prioritize one area over another. Often this takes the form of a short-term focus on exploitation rather than exploration. A third challenge to doing both simultaneously is that resources are limited. Providing more resources in one area means that the other area is less well resourced. Especially in situations where the need for either exploration or exploitation seems more pressing, the lure of prioritizing one over the other may become too great to resist. For example, many companies may be more inclined to improve on current operations, such as looking for cost-cutting opportunities during an economic downturn.

JOD, 1(2): 64-68 (2012)
DOI: 10.7146/jod.6344

Similarly, exploration may seem more appealing in a new or growing market segment or early in an industry's life cycle. A final challenge is that the structures and processes that promote exploration do not always facilitate exploitation. Recently, the CEO of 3M initiated Six Sigma management techniques across the organization, aimed at improving quality and controlling costs by reducing defects and variance in processes. While this exploitation-type effort worked well in many areas, it did not mesh well with the research and development function that is critical at 3M (Gunther, 2010).

How can managers design an organization that facilitates the pursuit of both exploration and exploitation? Ideally, how can you have an organization that continually refines what it is already doing while at the same time looking for promising new things to do? There is an ongoing debate in the scholarly community about how organizations can achieve these competing goals. Four main options have been suggested.

OPTION 1: OUTSOURCING

The first approach assumes that the tensions arising from trying to do both will make it too difficult to succeed. Instead, the organization focuses on one approach and outsources the other. For exploitation-focused firms, this means partnering or acquiring new technologies and products as the need arises. The Intel Capital fund, for example, provides the computer chipmaker the means to develop the market for new uses of its chips as well as provide an opportunity to acquire emerging technologies – both exploration-focused activities. Think also of pharmaceutical companies that in-license compounds that biotechnology firms have discovered or developed, while they focus on manufacturing, sales, and marketing activities. For exploration-oriented firms, this approach means focusing on the next big leaps in technology or product design while letting other companies handle other functions. Think here of many high-tech companies who partner with others to do contract manufacturing, leaving them to concentrate on product development. Indeed, you could view contract manufacturers as exploitation-focused firms who outsource exploration.

From an organization design standpoint, this approach is straightforward. You can design your organization with a clear focus. What this approach lacks is integration. It is critical to have a strong capability in managing the integration with the partner firms. For example, who chooses which new technologies to license or acquire? How are alliances managed within the organizational hierarchy? Which parts of the business are most dependent on these partners, and how integrated do they need to be? Is an acquisition kept autonomous or reconfigured within existing divisions? Without an integration capability, the success and failure traps mentioned earlier can lead to eventual obsolescence for exploitation-focused firms, or never profiting from the new options generated by an exploration-focused company.

OPTION 2: SEPARATE EXPLORATION AND EXPLOITATION, BUT ALIGN THEM

The need to integrate across the two types of activities shows up even more in the second option. Management recognizes the difficulty in doing exploration and exploitation simultaneously but takes the view that the best way to accomplish both is to develop an "ambidextrous" organization (O'Reilly & Tushman, 2004; *Organization Science*, 2009). Here the company creates organizational units with a clear focus on either exploration or exploitation, but rarely both. This type of organization is composed of "highly differentiated but weakly integrated subunits. While the exploratory units are small and decentralized, with loose cultures and processes, the exploitation units are larger and more centralized, with tight cultures and processes" (Benner & Tushman, 2003: 52). The required integration takes place largely at the leadership-team level, rather than being left to mid-level managers, since "the pressures on core business managers to meet current customer needs, optimize processes, and meet short-term financial expectations make it almost impossible for them to fully engage in exploring new opportunities at the same time" (Kates & Galbraith, 2007: 186).

The separation of the exploration-focused unit is often physical as well as organizational. Lockheed's own employees, for example, weren't told the location of the company's famed "Skunk Works" aerospace development center in the early days. Similarly, Steve Jobs famously moved the original Macintosh computer development team at Apple to a separate building and flew a pirate flag from the roof. The innovation-focused unit is kept away from the "corporate antibodies" that often do not see enough potential early in development to free up their own resources to support the new approach.

While this separation can protect innovation, it also makes the eventual required integration tougher to achieve. Since the reporting structures, goals, metrics, and rewards of the separate group may be very different, how do you build incentives and coordination mechanisms to integrate the different units? Xerox's Palo Alto Research Center provides a cautionary tale. PARC was located in California, far away from the corporate headquarters on the U.S. east coast. And while the technologies invented at PARC – such as laser printing, Ethernet, the graphical user interface, and the mouse – were critical to the technology industry, the fortunes made from them went to firms other than Xerox. Thus, one of the biggest challenges of ambidextrous organizations is how to leverage through exploitation what has been discovered by exploration – while safeguarding against knowledge spillover to rivals.

OPTION 3: CYCLE BACK AND FORTH BETWEEN EXPLORATION AND EXPLOITATION

The third option acknowledges that it can be a challenge to have both exploration and exploitation focused units in the same organization since these units will have different incentives, structures, and time horizons. Similarly, organization theorists recognize the benefits that arise from alignment around a singular focus. But since exploration or exploitation alone risks the success and failure traps, the third option is to cycle between these approaches. "Temporal cycling between long periods of exploitation and short bursts of exploration… [has] been identified as an alternative balancing mechanism that may be both logical and practical" (Gupta, Smith, & Shalley, 2006).

One of the most common areas of organization design where this shows up is in the centralization versus decentralization debate (Nickerson & Zenger, 2002), where each option offers benefits but a balance works best over time. The key idea "is that under certain conditions managers modulate between or among discrete structures to approximate, albeit temporarily, levels of functionality unachievable when organizations remain fixed with a particular structure" (Nickerson & Zenger, 2002: 548). Since this type of change can be disruptive and costly, scholars suggest that firms should pursue such a major change only after "a critical state of incongruence with the environment is reached" (Miller, 1982: 133).
The challenge here is twofold. Such a major reorientation requires many changes in the elements of organization design, such as structures, processes, coordination mechanisms, HR policies, metrics, and rewards. So getting the organization realigned around a very different approach is likely to be difficult. Second, for firms that reorient too regularly, employees may see the changes as faddish, and choose to simply wait out any reorientation rather than commit to it. Further, the need to reorient the firm when incongruence with the operating environment has been reached assumes that key decision makers are capable of sensing, evaluating, and responding to environmental signals.

OPTION 4: CONTINUOUS AND INCREMENTAL RECONFIGURATION

Since episodic and irregular change can be difficult to implement, the final option seeks to make change ongoing and incremental. Rather than involving the entire firm, the organization can adjust organizational units through continuous and incremental means. The idea here is that the boundaries between organizational units are redrawn, resources redeployed, and responsibilities reapportioned as needed. Eisenhardt and Brown (1999: 73) refer to this as an adaptive process of "patching," which is "…the adding, deleting, splitting, transferring or combining chunks of businesses." Beyond adaptation, the process of reconfiguring the firm may also serve as a mechanism for purposeful experimentation and the search for new opportunities (Karim, 2006).

This approach is the least well developed in the academic literature and thus the least well understood. It may be that this approach is a more frequent version of the third option (cycling back and forth). It is also possible that the evolution of reconfiguration occurs in one direction – from more exploration to more exploitation – as differentiated organizational units follow a product life cycle (Raisch, Birkinshaw, Probst, & Tushman, 2009). It can also be that different divisions, based on the turbulence of their markets, can each incrementally reconfigure to find the right balance between exploration and exploitation. Firms that patch and reconfigure are constantly updating their business units such that a unit is "small enough for agility and large enough for efficiency" (Eisenhardt & Brown, 1999: 74). A good example is Nokia, as it seeks to improve efficiency in its mobile phone division that serves the majority of developing markets, versus its mapping division that must remain flexible as it explores new ways to offer location-based services to its smartphone users. Thus, this approach is a more granular approach to the second option – exploration and exploitation may occur simultaneously but to differing degrees within different divisions.

Continuous, incremental redesign may sound ideal. However, frequent reconfiguration of organizational boundaries can be costly and not necessarily successful. Unit reconfiguration announcements – where most units are reconfigured either to increase efficiency or to pursue growth opportunities in new product markets – initially increase shareholder wealth but then face a period of decline in earnings performance (Brickley & Van Drunen, 1990). Not only may the processes (of redrawing, redeploying, and reconfiguring) be costly to accomplish, but they also may be disruptive to the organization if not integrated seamlessly into the existing structure. This is a micro version of the challenges associated with option one but occurring within a single firm. Scholars have found that the art of frequent, incremental reconfiguration is also something that must be learned over multiple experiences before leading to subsequent innovations (Karim, 2009). Thus, firms that have limited resources for reconfiguration may be dissuaded from this type of process.

CONCLUSION

Note that there is wide variety in the options discussed above. The first option takes the view that maintaining a balance within the same organization is close to impossible. The last option holds that it is not only possible but advisable to try to do both. The other two options are somewhere in the middle. The lack of consensus reflects the significant challenge that exploration and exploitation present for any organization.

Following contingency logic, the right choice of organization design always "depends" – on features of the organization, its strategy, its operating environment, and so on (Burton, DeSanctis, & Obel, 2006). Thus, factors both internal and external to the firm should guide the choice of option. After all, an organization with little experience in selecting and managing alliance partners would likely have a difficult time taking the first approach. Likewise, the capabilities related to managing ongoing organizational change would be critical to the last approach. While each of the options has its challenges, not recognizing and attempting to deal with the fundamental challenge of doing both exploration and exploitation would be the worst response.

REFERENCES

Benner M, Tushman, ML. 2003. Exploitation, exploration, and process management: the productivity dilemma revisited. *Academy of Management Review* 28(2): 238–256.

Brickley JA, Van Drunen, LD. 1990. Internal corporate restructuring: an empirical analysis. *Journal of Accounting and Economics* 12(1-3): 251–280.

Burton RM, DeSanctis G, Obel B. 2006. *Organization Design: A Step-by-Step Approach*. Cambridge University Press, New York.

Eisenhardt KM, Brown SL. 1999. Patching: restitching business portfolios in dynamic markets. *Harvard Business Review* 77(3): 71–82.

Gunther M. 2010. 3M's innovation revival. *Fortune* (September 24): 73–76.

Gupta AK, Smith KG, Shalley CE. 2006. The interplay between exploration and exploitation. *Academy of Management Journal* 49(4): 693–706.

Karim S. 2006. Modularity in organizational structure: the reconfiguration of internally developed and acquired business units. *Strategic Management Journal* 27(9): 799–823.

Karim S. 2009. Business unit reorganization and innovation in new product markets. *Management Science* 55: 1237–1254.

Kates A, Galbraith JR. 2007. Designing Your Organization. Jossey-Bass, San Francisco.

March JG. 1991. Exploration and exploitation in organizational learning. *Organization Science* 2(2): 71–87.

Miller D. 1982. Evolution and revolution: a quantum view of structural change in organizations. *Journal of Management Studies* 19(2): 131–151.

Nickerson JA, Zenger TR. 2002. Being efficiently fickle: a dynamic theory of organizational choice. *Organization Science* 13(5): 547–566.

Organization Science. 2009. Special Issue: Organizational Ambidexterity. 20(4): 685–828.

O'Reilly CA, Tushman ML. 2004. The ambidextrous organization. *Harvard Business Review* 82(4) 74–81.

Raisch S, Birkinshaw J, Probst G, Tushman, ML. 2009. Organizational ambidexterity: balancing exploitation and exploration for sustained performance. *Organization Science* 20(4): 685–695.

TIMOTHY N. CARROLL

Assistant Professor
Moore School of Business, University of South Carolina
E-mail: carroll@moore.sc.edu

Acknowledgements: My thanks to Samina Karim and Charles Snow for comments on earlier versions of the paper.

DESIGNING THE FIRM TO FIT THE FUTURE

RAYMOND E. MILES • LAURENT SCARINGELLA

Abstract: Most firms identify market opportunities for their new technologies after they have been developed. This article discusses the design of a "futures group" which can help to synchronize a firm's technology and market development. A futures group designed to span more than one organization could lead to simultaneous market development for multiple technologies.

Keywords: Organizational alignment; organizational fit; technology commercialization; market intelligence

In the world of design, form is expected to follow function. An architect designs a building to meet a specific purpose, and an aeronautical engineer designs an aircraft to meet specific flight requirements. Similarly, firms are designed to produce goods or services for a particular market with organizational features and costs specified in their business plans. Thus, it follows that how the organization's resources and guidance mechanisms are arranged should draw early and detailed managerial attention. Logic, however, may not prevail, and the design of the structures essential to house the organization's resources, and the processes necessary to direct and control them, tend to become the focus of management attention only after problems arise. Moreover, in today's increasingly competitive global economy, design flaws will become more costly and belated managerial attention more injurious as the pace of innovation accelerates, requiring firms to rapidly assemble and apply complex technical and market knowledge. Indeed, it appears that firms in many industries are entering a period in which they will need organizational designs that help them *anticipate* rather than follow technological and market developments.

In this article, we discuss a common organizational problem: Although market intelligence regarding potential technology commercialization opportunities is clearly valuable (Frishammar, Lichtenthaler, & Rundquist, 2012), firms are not equally adept at guiding their innovation efforts with both technical and market knowledge (Burgers, Van Den Bosch, & Volberda, 2008). As a result, lengthy and costly time gaps may occur before a firm can find market applications for its technological innovations. Building on our ongoing research, as well as several examples of firms which have attempted to solve this problem, we describe the essential features of an entrepreneurially oriented "futures group," an organizational mechanism that firms can use to explore and develop future markets for their new and emerging technologies.

A RECURRING CHALLENGE

In organizations, the interaction of function and form is a recurring challenge. Over the past two centuries, the development of new technologies has stimulated the growth of new industries populated by new firms bringing new products and services to an expanding set of markets. During the co-evolution of technologies and markets, new capabilities have been utilized in new firm processes organized and managed in new ways. In recent decades, attention has been directed at economic sectors such as biotechnology and information processing where research draws on multiple scientific knowledge bases to create a large and growing body of technical knowledge awaiting full application in innovative products

 Journal of Organization Design
JOD, 1(2): 69-74 (2012)
DOI: 10.7146/jod.6343

and services.

As knowledge-driven product/service complexity increases, individual firms often have difficulty envisioning the full market potential of emerging technological innovations. Although applications in existing markets may be foreseeable, applications in adjacent markets typically are not. Therefore, new firm and even inter-firm strategy-structure designs need to be developed if the global economy is to take full advantage of expanding knowledge to drive product and service innovations.

Several years ago, Miles, Miles, & Snow (2005) speculated about the shape of organizational strategies, structures, and processes capable of producing major increases in knowledge utilization, suggesting that future organizational designs should enable the formation of collaborative communities of firms in which multiple complementary technologies are combined to produce a continuous stream of new products for existing and related markets. Such designs would allow the formation of collaborative ventures across markets and technologies – ventures that would broaden the technology and market vision of community member firms and stimulate further innovations by combining knowledge across firms' business models and customer markets. However, while one successful collaborative community of firms has been studied (Snow, Fjeldstad, Lettl, & Miles, 2011), it focuses on only a single technology with flexible employability (the blade processor technology in the computer server industry). Although experiments with new community-based designs are underway, the development of a multi-technology collaborative community of firms may be years away. It is, therefore, much easier to imagine how individual firms can move toward innovation-facilitating organizational processes, designs which will inevitably include at least some inter-firm collaboration involving related technologies, products, and/or markets.

The overall challenge is to design a firm in a manner that enhances its entrepreneurial capabilities. Entrepreneurs typically create wealth by adapting existing technologies or products to serve new purposes in new markets, a process that is facilitated by inter-firm collaboration across complementary markets. Entrepreneurial innovation of this sort, however, violates the logic built into most organizational designs, a logic that links existing technical knowledge with existing product or service lines. How can a firm improve its ability to expand into new markets, either by adapting its technologies or by linking to related technologies? And how can it speed up the process of doing so?

FUTURE FIT DESIGN EXPERIMENTS

Technical and market knowledge can be infused directly into firms through government action (Link & Siegel, 2007). In contrast to the free-market approaches of the US and the UK, where government-supported research finds its way entrepreneurially into firms clustered around major research universities, both Taiwan and the People's Republic of China have created governmental agencies that offer technical and market assistance directly to firms. The Taiwanese approach evaluates technologies developed elsewhere for possible matches with Taiwanese firms (Mathews, 1997; Mathews & Cho, 2000). Agency employees are even free to form their own firms to take advantage of promising technologies. This approach produced a new player (Acer) in the global computer industry in an economy functioning mostly as a components supplier and original equipment manufacturer for firms in other countries (Mathews & Snow, 1998). The newer and massive PRC investments are less easily evaluated at this point but are expected to have major long-term impact across several industries in the global economy.

A recent experiment in France aimed at enhancing the flow of technical knowledge across firms in the Grenoble region falls between the planned government investments in China and the entrepreneurial university-firm knowledge flows found in the US and UK. CEA-LETI, a redesigned version of several government agencies originally created to stimulate both military and civilian applications of atomic energy, is now charged with providing research knowledge to nanotechnology, computer chip, and other high-tech firms in the Grenoble region (Scaringella & Miles, 2011). CEA-LETI performs and provides research to firms for a fee, and it can create its own start-up enterprises in new product areas and markets. To date, CEA-LETI has created over 40 high-tech start-up firms, many of which have been acquired by larger firms. In addition to CEA-LETI, the Grenoble-based CEA agency has created two new units within the past eight years, one focused on a cluster of software firms close to Paris and the other aimed at supporting the flow of technical knowledge into new energy systems for housing and transportation.

In the private sector, one company that has successfully wrestled with the problem of how to organize to explore new businesses while operating mature businesses is IBM (O'Reilly, Harreld, & Tushman, 2009). After an internal company analysis revealed six major reasons IBM routinely missed new technology and market opportunities, the firm developed the Emerging Business Organization (EBO) initiative in 2000. The elaborate EBO process systematically explores, creates, and tests new business units that are then either grown or terminated. In less than ten years, 25 EBOs were launched. Three failed and were closed, but the remaining 22 now produce more than 15% of IBM's revenue. Of course, IBM is a huge corporation with enough internal resources to explore and experiment. Smaller, less well-endowed firms need organizational units that can emulate EBO's entrepreneurial process without overwhelming their available resources.

Lastly, an important component of a "future fit" organizational model designed to push new technologies beyond current market uses, and in the process suggest directions for further technological development, may be emerging in the wireless communications industry where firms such as Apple have encouraged and facilitated communities of designers who create applications that extend the information seeking and processing uses of the core technology. Knowledge flows within such communities combine market and technical knowledge in a most entrepreneurial fashion. Indeed, the wide array of available applications represents an ongoing exploration of consumer needs and desires, with usage data pointing the way to future technological and market developments.

Both the public and private sector efforts discussed above suggest mechanisms that could be incorporated into the design of an organizational unit focused on the continuous and simultaneous development of technologies and markets. The purpose of such a unit would be to integrate technology development, market expansion, and the venture capital needed for research, experimentation, and capability development.

A POSSIBLE FUTURE FIT DESIGN

One can generally imagine an organizational unit that combines emerging technical and market knowledge to guide next-stage product/service innovations, a unit that contains many of the features described in the examples above. The challenge is to design the structure and the managerial processes that would assemble, maintain, and direct such a unit within the resource constraints of a typical firm. Assembling the R&D or technological component seems the lesser challenge, as one can simply extract technical specialists from the firm's existing R&D units and assign them to the firm's futures group. The key design issue for this segment is how to arrange those technical specialists so that they can both share the skills and knowledge that define their disciplines and explore knowledge combinations and applications outside their normal uses. The solution is to develop a matrix-like design that simultaneously groups specialists by discipline and by current and potential markets. Creating the matrix axis focused on current and related markets seems straightforward, assuming that the firm is willing to search inside the organization and out for entrepreneurially inclined individuals and to reach across into complementary firms in what we imagine will be an ever-broadening set of related or potentially relatable markets. For example, the French agency's knowledge transfer unit that addresses the software industry is arranged as a matrix while the very newest unit focused on energy systems is organized as a single adaptive team structure.

A large initial financial investment is required for a futures group, an investment that encourages and empowers a start-up period of knowledge sharing not only across markets but across scientific fields as well. A futures group would focus on first one, and then another, nearby or more distant market segment, looking at both the technical knowledge employed and the product/service innovation stream and market response over recent years. Such broad discussions seem likely to elicit a beginning stream and then a rush of suggestions for expanding current product/service innovation efforts and exploring possible new technical and market domains.

The key resource allocation decision, we predict, will quickly turn from encouraging interest to forming and launching innovation projects. Resource allocation and guidance in the futures group are provided by the entrepreneurial axis of the matrix, with each entrepreneur/ manager free to launch one or more market-serving ventures using an initial pool of funds to buy the time of the needed technical talent. The exact amount of funds provided initially to each entrepreneur will vary by firm size and current resources, but a portion of the funds is already available in the budgeted salaries of the technical specialists and the entrepreneurs recruited from existing departments. As projects move forward, additional funding will be required. This suggests the need for an in-house venture capital fund and allocation committee, a process that a number of firms, notably Hewlett Packard, have used successfully in the past. More recently, Shah, Ortt, and Scholten (2010) have described venture development mechanisms in three large firms noted for their innovation efforts: Royal Dutch Shell, Nokia, and IBM. The Shell approach encourages bottom-up ideas that are evaluated and funded by a team of upper-level managers and directed toward either internal development or external venturing with selected partners. Nokia uses similar mechanisms to encourage both internal development of new technologies and external ventures across a wide group of affiliates. Lastly, IBM's exploration efforts, as illustrated by the EBO initiative described above, are guided by top-down decisions, but the development of market applications usually involves a variety of partner firms.

Clearly, a wave of experiments with new product /service designs outside the firm's existing innovation stream requires coordination mechanisms and collaboration capabilities. One key mechanism we have suggested is an "idea bank" into which each entrepreneurial venture is entered and updated as progress occurs (Miles et al., 2005). Such a data source will quickly reveal both potential paths for collaborative action across project teams and potential duplications. Duplications can and probably should be contained by the venture capital decision-making process, and collaborative developments can be facilitated by top management encouragement and support. It is our view that a futures group will both attract and expand collaboratively inclined technical and market-focused talent and that managerial endorsement, reinforced by venture funding decisions, will guide the group toward further collaborative behaviors. In addition, reward structures within the futures unit will need to be designed to reinforce collaboration across the entire set of innovation projects undertaken.

Once product/service prototypes begin to emerge across the firm's current and related markets, next-stage procedures must be in place. A single firm can handle only a few of the entrepreneurial ventures one can imagine emerging from the futures group. Two potential routes to market success are common across high-tech firm clusters: the spin-off venture and a collaborative multi-firm alliance. The futures group, assisted by the venture capital committee, can approve a limited number of spin-offs, with an equity position taken by the mother firm and remaining funding coming from outside venture capital investments that validate the spin-off's entrepreneurial promise. This is an internal version of the start-up activities undertaken by the French agency described earlier and is a common practice in the Shell, IBM, and Nokia examples.

A second route to market is a collaborative venture with one or more firms in complementary markets. A given firm's futures group, with entrepreneurial talent drawn in part from complementary markets, may well have potential collaborative relationships built into it, but such firm-to-firm linkages are dependent on the exploring firm's demonstrated commitment to equitable treatment of potential partners. The conditions essential for both internal and external collaborative relationships are described in Miles et al. (2005). A firm that demonstrates its trustworthiness in its initial collaborative ventures is likely to find numerous potential partners for its future collaborative ventures. In our view, collaborative entrepreneurship is the most powerful mechanism for generating multi-disciplinary knowledge-driven innovations. Moreover, one firm's successful futures group could inspire similar units across firms in complementary markets, units that are inclined from the beginning toward initiating and responding positively to inter-firm collaborative initiatives.

If a firm's futures group flourishes, it will be vulnerable to both the up-side and the down-side challenges of organizational success. The up-side challenge is how to maintain and grow its own resources to prevent entrepreneurial stars from either jumping to other firms or starting up their own ventures, likely taking with them members of their project teams. The down-side challenge is how to manage possibly deteriorating relationships with the firm's less entrepreneurially oriented units which have generated the venture capital that the futures group employed on its way to success. Indeed, the mind-set and stylistic differences between a futures group and established units focused on current products and markets mirrors the continuing challenges inside most firms created by the interaction of R&D and marketing. R&D, if it does its appointed job, is focused on evolving science and its longer-term implications, while marketing is focused on near-term success and failure across existing markets. Clearly, these are predictable challenges that require well-designed intra-firm collaborative mechanisms. Also, firms must consider how reward systems will evolve as entrepreneurial ventures succeed, including designs that increase the managerial and financial rewards flowing across both the forward-looking and stable segments of the firm to reflect their joint contributions to firm performance. Successful firms, it appears to us, become proficient in combining the capabilities to exploit existing technologies within current and related markets with the capabilities to explore related technologies and their potential market applications.

One approach to dealing with the challenges of success is to pose them in advance to the leadership teams occupying the entrepreneurial axis of the futures group's matrix. One of the skills of entrepreneurs is bridge building, a useful orientation and skill if one is inclined toward serial entrepreneurial activities. Ideally, an internal entrepreneurial unit could become an initiating agency in the creation of both internal and multi-firm innovation communities.

CONCLUSION

The global economy has entered a period in which scientific and technical knowledge is becoming more complex and faster flowing. Taking full advantage of this dynamic resource represents a potential competitive advantage to firms and a potential comparative advantage to national and regional economies. The challenge to the firm is to exploit its current market-focused technology while simultaneously exploring the full range of market opportunities that its technical knowledge might engage. What we know is that most existing organizational arrangements, which tend to separate knowledge generation from knowledge application, result in only a fraction of available knowledge finding profitable use.

In our view, organizational design experiments aimed at capturing all of the market opportunities presented by new and combined technologies are clearly warranted. We have suggested one such experiment in the form of a futures group. Entrepreneurial units of this kind may be threatening to established organizational units and, if attempted, even more challenging if they succeed. It is, however, just this sort of design challenge that successful firms point to as crucial turning points in their efforts to become more innovative. Based on our continuing research, we believe that entrepreneurial talent and motivation are far more abundant than most organizational processes recognize and use. The design and implementation of a futures group aimed precisely at giving such skills and interest both the freedom and the support to flourish seems to be a worthy objective.

REFERENCES

Burgers JH, Van Den Bosch FAJ, Volberda HW. 2008. Why new business development projects fail: coping with the differences of technological versus market knowledge. *Long Range Planning* 41: 55–73.

Frishammar J, Lichtenthaler U, Rundquist J. 2012. Identifying technology commercialization opportunities: the importance of integrating product development knowledge. *Journal of Product Innovation Management* 29(4): 573-589.

Link AN, Siegel DS. 2007. *Innovation, Entrepreneurship, and Technological Change.* Oxford University Press, Oxford, UK.

Mathews JA. 1997. A Silicon Valley of the east: creating Taiwan's semiconductor industry. *California Management Review* 39(4): 26–54.

Mathews JA, Cho DS. 2000. *Tiger Technology: The Creation of a Semiconductor Industry in East Asia*. Cambridge University Press, Cambridge, UK.
Mathews JA, Snow CC. 1998. A conversation with the Acer Group's Stan Shih on global strategy and management. *Organizational Dynamics* 25: 65-74.
Miles RE, Miles G, Snow CC. 2005. *Collaborative Entrepreneurship: How Communities of Networked Firms Use Continuous Innovation to Create Economic Wealth*. Stanford University Press, Stanford, CA.
O'Reilly CA, Harreld JB, Tushman ML. 2009. Organizational ambidexterity: IBM and emerging business opportunities. *California Management Review* 51(4): 75–99.
Scaringella L, Miles RE. 2011. Exploring the paths of firms' successes and failures in high technology markets. Working Paper, Haas School of Business, University of California, Berkeley.
Shah CM, Ortt JR, Scholten V. 2010. Building a radical innovation mechanism at large firms. In L. Al-Hakim and C. Jin (Eds.), *Innovation in Business and Enterprise: Technologies and Frameworks*: 120–134. IGI Global, Hershey, PA.
Snow CC, Fjeldstad ØD, Lettl C, Miles RE. 2011. Organizing continuous product development and commercialization: the collaborative community of firms model. *Journal of Product Innovation Management* 28: 3–16.

RAYMOND E. MILES

Professor Emeritus
Haas School of Business, University of California, Berkeley
E-mail: miles@haas.berkeley.edu

LAURENT SCARINGELLA

Assistant Professor
ESC Rennes School of Business, France
E-mail: laurent.scaringella@esc-rennes.fr

DISOBEYING POWER-LAWS

PERILS FOR THEORY AND METHOD

G. CHRISTOPHER CRAWFORD

Abstract: The "norm of normality" is a myth that organization design scholars should believe only at their peril. In contrast to the normal (bell-shaped) distribution with independent observations and linear relationships assumed by Gaussian statistics, research shows that nearly every input and outcome in organizational domains is power-law (Pareto) distributed. These highly skewed distributions exhibit unstable means, unlimited variance, underlying interdependence, and extreme outcomes that disproportionally influence the entire system, making Gaussian methods and assumptions largely invalid. By developing more focused research designs and using methods that assume interdependence and potentially nonlinear relationships, organization design scholars can develop theories that more closely depict empirical reality and provide more useful insights to practitioners and other stakeholders.

Keywords: Power-law distributions; Gaussian statistics; Pareto; nonlinear statistical methods; theory building

As myriad studies in nearly every area of business and science indicate, the norm of normality is a myth. Instead, we exist in a world where power-law (i.e., Pareto) distributions are ubiquitous. In contrast to the traditionally assumed normal (Gaussian) distribution of organizational outcomes, where events are completely independent and identically distributed, power laws identify the fundamental interconnectedness and interdependence of events (Andriani & McKelvey, 2009). As shown in Figure 1, power-law distributions are highly skewed, with long, fat tails (a downward-sloping straight line when plotted on log-log axes) that identify outliers (i.e., extreme events). When graphed on regular scales, the power-law distribution looks like Figure 1a; log-log axes are shown in Figure 1b. These distributions are interesting because of the various outcomes contained in their elongated tail, represented by the "Paretian world" area to the right of the shaded region in Figure 1b.

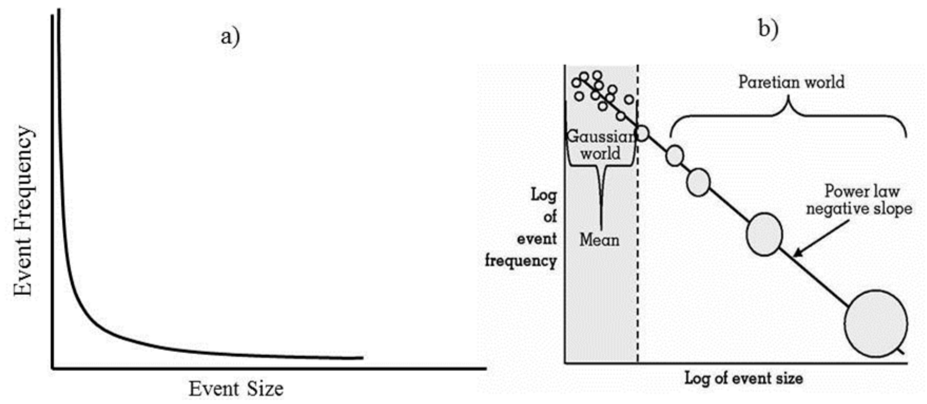

Fig. 1. Power-Law Distributions: Normal and Log-Log
Source: Figure 1b is from Boisot and McKelvey (2010).

The shaded Gaussian region on the left of Figure 1b represents the vast majority of organizations. There is an obvious contrast in size and potential scope of influence as firms move down the slope into the Pareto region. Though infrequent, an outcome in the long

DOI: 10.7146/jod.6419

tail of a power-law distribution – the large circle at the bottom right of Figure 1b – is of disproportionate influence on the entire system. To illustrate, the small circles in the upper left might represent the 17 million Mom & Pop retail stores in the United States, while the largest circle in the bottom right might represent Walmart. Accordingly, this distribution represents something qualitatively different that must be taken into account. The distinctiveness of the distribution and its complex effects make it relevant to both practitioners and scholars. Empirical studies have discovered power-law patterns in nearly every aspect of the internal and external contingencies explored by organization design scholars. For example, such distributions have been found in U.S. firm size, overall industry structure, competitive performance advantages, industry sector and firm growth rates, firm survival and exit, network structure, market share prices, product innovations and technological breakthroughs, entrepreneurial growth expectations, new venture performance, and individual performance – and in many others in physical, natural, biological, and social systems.

Power-law distributions emerge as a result of tension and connectivity dynamics among agents in a system. However, it is often difficult to see – let alone understand – these patterns if they are viewed only at one level. Although power-law distributions are pervasive in domains relevant to organization design, the importance of these unique statistical signatures is seldom explored.

POWER-LAW EFFECTS ON EMPIRICAL OBSERVATIONS

Traditional (Gaussian) statistical analyses are not applicable to firms in all regions of the distribution when power laws are present (Boisot & McKelvey, 2010). Though empirical reality continually displays evidence of skewed outcomes, scholars continue to use Gaussian statistical techniques that assume normal distributions of outcomes, linear relationships among variables, stable means, finite variance, and independence of events. Indeed, using Gaussian assumptions and methods to explain power-law phenomena can lead to inaccurate conclusions, under-specified theoretical models, and misleading normative recommendations, all of which reduce the credibility of scholarly research (O'Boyle & Aguinis, 2012). The prevalence of power laws has implications for both research design and statistical analysis.

Research Design Implications

Starbuck and Nystrom (1981) suggested decades ago that useful prescriptions for organization design do not come from conducting indiscriminant empirical studies. Instead, they argued, "truly innovative designs have to originate in deviant cases or in fantasies rather than in statistical norms (1981: 8)." Here, "deviant" cases are those outside "normal" (i.e., in the tail of the power-law distribution) because of their size, rarity, and potential scope of influence on the environment. Several aspects of power-law distributions will highlight their applicability to the empirical study of organization design. As the Gaussian and Paretian regions of Figure 1 suggest, there are distinct differences in the inputs and outcomes for firms in each region. First, these distributions exhibit data that are both linear and nonlinear. While assumptions of independence and additive relationships can aid in understanding linear relationships, these assumptions are wholly violated in nonlinear relationships. This also suggests that studying extreme, nonlinear outcomes – the companies that organizational scholars frequently use as deviant case examples and seek to explain such as Apple, Facebook, and Enron – is implicitly hindered by the use of traditional statistical methods. Second, once firms are large enough to be in the Pareto region of the distribution, their activities become multiplicative and nonlinear, where they have the potential to influence the outcomes of other firms in the sector.

When we use Gaussian methods to study firms in a population, we are not measuring performance as much as we are constraining it. Few phenomena adhere to a power law over all values. Instead, the power law most often applies for values greater than some minimum – this is the tipping point of the distribution, where the tail begins. As an example, the graph in Figure 2 shows an analysis of annual revenue of the Inc. 5000 Fastest-Growing Companies in the United States. Using the plfit.m script from Clauset, Shalizi, and Newman (2009) in MATLAB to construct a semi-parametric bootstrap maximum likelihood estimation of fit with a power-law model, the graph shows the data tip from linear to nonlinear when firms reach $158M in revenue. At this size, firms have a much greater potential to influence the surrounding environment (i.e., to have co-evolutionary effects). Therefore, it is important to design studies that include entire populations of interest – or at least large random samples thereof in order to maximize data collection in all regions of the distribution.

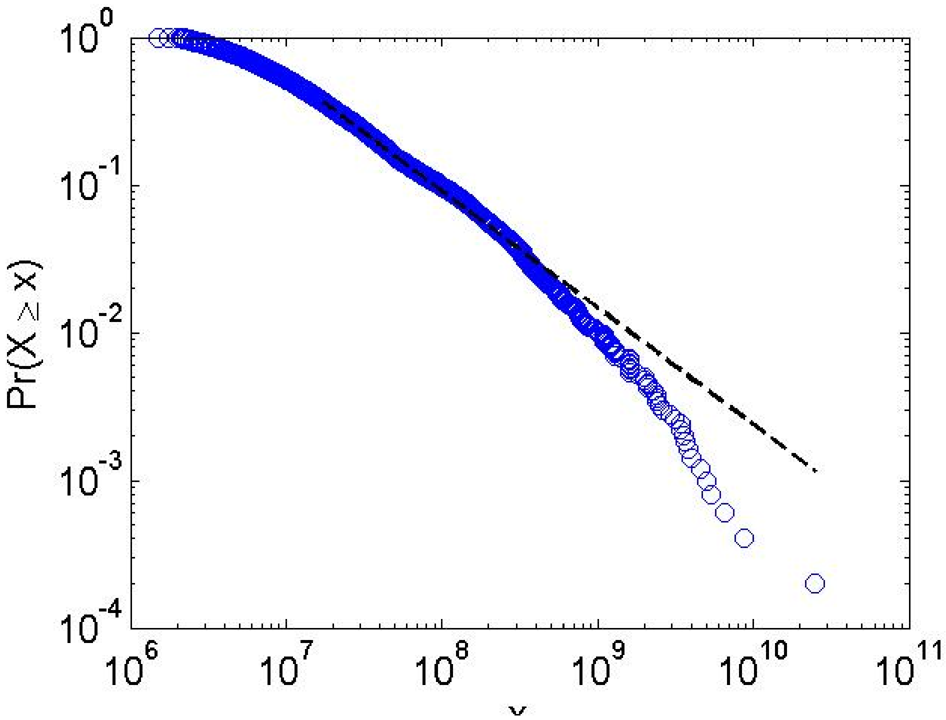

Fig. 2. Maximum Likelihood Estimation of Power-Law Fit: Inc. 5000 Fastest-Growing Private U.S. Companies
Note: The entire dotted line represents the maximum likelihood estimate of the power-law tail's slope, calculated as 1.79. The tipping point of this distribution, where nonlinear and co-evolutionary effects begin, is $158M. Kolmogorov-Smirnov goodness of fit is 0.036 (≤ 0.10 is desirable).

Statistical Implications

If viewed through the lens of traditional statistics, entities in the tail of the distribution are outliers – they are either viewed as random or a hindrance to obtaining statistical significance. However, rather than being deleted or transformed, these outliers often should be the ones that really matter to organization design scholars. The statistical average does not help scholars understand the true dynamics of the environments in which those firms exist, and it does not provide much instructive relevance for managers. What matters are the extremes! When data are skewed, especially where agent connectivity and interdependence are prevalent, it is likely that power-law dynamics are influencing the distribution. As an example, a normal distribution has a skewness of 0, and data are considered skewed if that number is above 3 (Greene, 2007). Interestingly, in Figure 2, the skewness of these data is 36. Thus, as skew increases, the more the distribution has the potential to exhibit power-law characteristics. All of this becomes very problematic for researchers using Gaussian techniques that assume normal distributions with stable means and finite variance, a problem that has been extensively documented (Andriani & McKelvey, 2009; O'Boyle & Aguinis, 2012; Simon, 1955).

In power-law distributions, the mean is unstable and variance is nearly infinite; therefore, no single observation can represent the average of the system. Revisit Figure 1b: Whereas the tip of the downward-pointing bracket is implied as the mean of the distribution, it is probably closer to the median. In highly skewed distributions, extreme values on the right often pull the mean beyond the lower bound of the power-law tail. According to O'Boyle and Aguinis (2012), this suggests that nearly 70 percent of the population is performing below average. As in Figure 2, the mean of the distribution is $73M while the median is only $11M. This implies that any explanatory or predictive theory built or tested using linear statistical methodology on the decidedly non-normal inputs and outcomes at every level in the domain must include a discussion about how the violation of Gaussian assumptions affects the analysis.

The most important variance in the data is buried by traditional robustness techniques. To wit, a common rebuttal from econometric scholars is, "Just transform the data to reduce the influence of outliers." Doing so undoubtedly increases the probability of a significant finding (and may increase the probability of a favorable review by editors). However, such a transformation of data obfuscates the actual effect of the differences that firms experience. In a study of the retail industry, for example, transforming the data does not reduce the real magnitude of Walmart's influence on a corner Mom & Pop store. Thus, deleting or manipulating outliers to achieve statistical significance is misguided. Schoonhoven (1981) suggests that assuming linear relationships is "misplaced" and that testing for nonlinear effects should be mandatory in all empirical analyses.

It would behoove scholars to more thoroughly understand their data prior to conducting Gaussian analyses. Two questions organization design scholars can ask as they develop their research projects are: "Can the performance of an outlier influence the outcomes of others in the distribution?" and "Can one node accurately represent the average of the population?" If the answer to the first question is "yes" or the second question "no", then quantitative nonlinear techniques like Poisson processes, Bayesian neural networks, historical extreme event analysis, or deep structure analysis may produce more accurate descriptions of the system outputs (O'Boyle & Aguinis, 2012). Similarly, techniques using non- or semi-parametric distributions, or agent-based models, where probabilistic interactions among firms can be simulated in a virtual environment, could be used to more accurately reflect empirical reality.

POWER-LAW EFFECTS ON THEORY BUILDING

When it is likely that empirical observations are influenced by power-law effects, theory building efforts should reflect their presence. Power laws are called "scale-free" distributions because they look the same regardless of the scale used to measure them. In these distributions, the relationships among the size of the events are fractal – they have self-similar behavioral patterns and physical characteristics, where the small appears similar to the big, and individual sub-parts look the same as the whole (West & Deering, 1995). Here, theory development is demanding because it requires knowledge about the whole system and about the underlying emergence of all the sub-systems.

Researchers must investigate causality from both the bottom up and the top down. Simon (1968) and others suggest that the emergence of power-law distributions is driven from the bottom up by the simple rules of the agents (e.g., workers, firms, teams) as they interact within a system. From the bottom up, rules represent an agent's recursive decision-making heuristics for achieving desired outcomes. These heuristics influence an agent's habitual behavioral strategies for interacting with the environment. Over time, as successful strategies are given positive feedback from the environment, power-law patterns of outcomes emerge. As agents accumulate resources (e.g., revenue, employees) and become large enough to be in the tail of a distribution in a local environment, there is an increased potential for co-evolutionary effects on the global environment, one level of analysis higher. From the top down, rules (e.g., corporate goals, cultural context, regulatory restrictions) and inputs (e.g., quality and quantity of both competitors and resources in the environment) impose tension on the sub-systems. Together, the rules and inputs throughout the entire system require theory that explains this co-evolving causality. Thus, power laws may be generated in a process as shown in Figure 3.

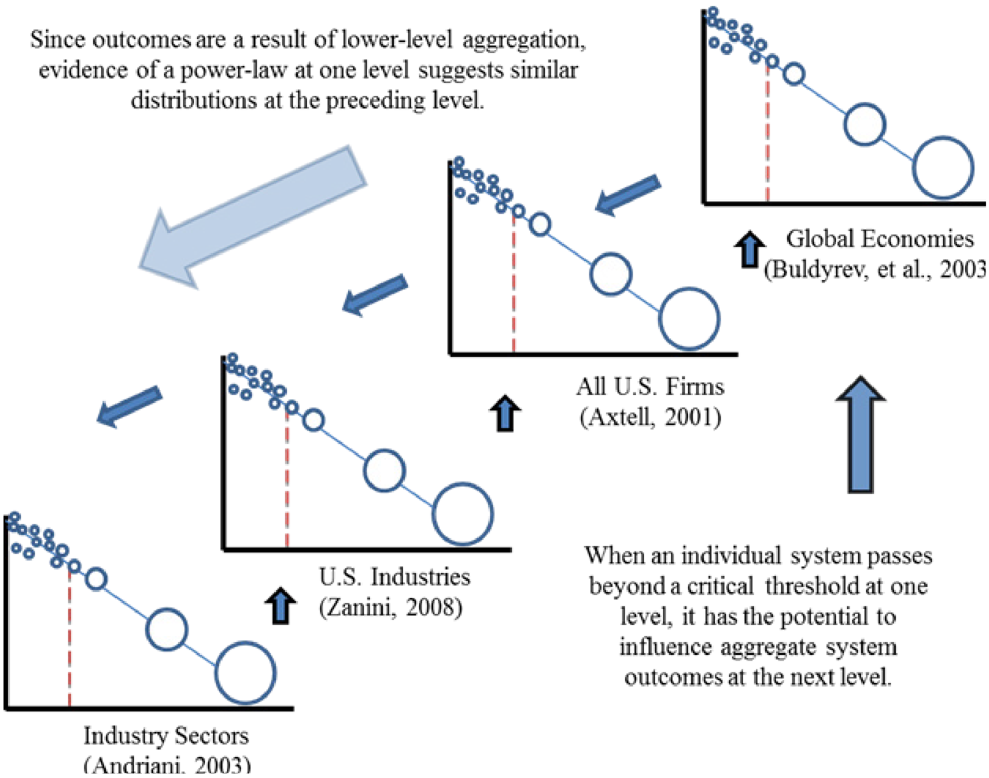

Fig. 3. Cycle of Bottom-up Emergence and Co-evolutionary Causality

When power-laws are present, theories to explain them are scale-free. In this case, an explanation for outcomes at one level of analysis should also explain inputs and outcomes at preceding levels. A scale-free hypothesis can pose as a simple theory that provides the best scalable explanation for the empirical regularities found in the data – one that seeks inference toward the best explanation by examining a mass of data and suggesting a plausible explanation for the patterns. Such a hypothesis is a simple, parsimonious, plausible, and falsifiable theory (Popper, 1963) for the emergence of power-law distributions and for the extreme outcomes therein.

In contrast to the probabilistic determinism of traditional statistics, a scale-free hypothesis needs to identify inputs and rules that could connect to produce an extreme outcome. A scale-free theory provides plausible anticipation rather than prediction. What is "plausible"? As Simon (1968:449) explains, "It is not inconsistent with everyday knowledge. At the moment they [i.e., the simple set of generative mechanisms] are introduced, they are already known (or strongly suspected) to be not far from the truth." This may create push-back, though, from reviewers who have experience and comfort with highly refined econometric models that supposedly control for alternative configurations and provide robust statistical significance. Thus, scholars can improve the efficacy of their theories and methods by integrating power-law logic into their research designs to facilitate community-wide acceptance.

INCORPORATING POWER-LAW REASONING INTO RESEARCH DESIGNS

Scale-free theory building efforts need to integrate disparate empirical knowledge sources. The more the knowledge is representative of the entire population – whether from random-digit dialing, meta-analysis, or experimental studies – the better. Subsequent empirical testing will require nonlinear methods to discover generative mechanisms. For example, Crawford (2012) hypothesized that a new venture founder's resource endowments and expectations for future growth generated the highly skewed distribution of outcomes in entrepreneurship. Analyzing organizational forms in three representative samples, at three different stages, at multiple time periods – starting with nascent pre-organizing expectations, to emergent outcomes, to active start-up firms, to hyper-growth companies (N = 11,000) – the study used semi- and non-parametric bootstrap simulations to find statistically identical power-law slopes for nascent expected growth and for firm growth rate in all three samples. This suggests that expectations for growth influence the distribution of outcomes throughout the domain. Future theory-building efforts in this subject area should account for (and collect data on) growth expectations – from both the top down and the bottom up – at multiple levels to limit unobserved variable bias (Carroll & Harrison, 1998).

Scholars need to study the causal dynamics that underlie the emergence of power-law distributions by linking methodologies to the points on the distribution where they are most useful. Scholars may hypothesize that power laws come from "one generative mechanism" as suggested by Boisot and McKelvey (2010) or from "a simple set of mechanisms" as posited by Simon (1968). Depending on the research question of interest, scholarly investigation to explain power laws should be organized around one definition of generative mechanisms or the other. Interest in one mechanism should investigate the agent rules for interaction at multiple levels of analysis and points in time, searching for a common dynamic at the beginning of the focal process that continues to the final outcomes of interest – this will be indicated by a universal slope, α, of the power-law tail. Interest in a set of mechanisms should identify power laws at multiple units and levels of analysis, focusing on both top-down and bottom-up rules and inputs like the type discussed in the previous section – for each distribution, significance can be tested with Komolgorov-Smirnov statistics or p-values. Similarly, studying extreme outcomes in the tail of the distribution with multiple qualitative field studies or a hermeneutics study can provide insight and theoretical grounding.

Boisot and McKelvey (2010:428) maintain that "research must engage with the power-law distribution as a whole, without privileging one particular region at the expense of another." This may not necessarily be required, however. Higher-level (i.e., mid-range) theory building must account for the emergence of extreme outcomes in the domain, but instead of amassing all outcomes together, a lower-level theory may focus on one particular region of the distribution. It is especially important to specifically state any theoretical assumptions and boundary conditions that apply to different treatments of the data in different regions of the distribution. Either way, scholars must start with the interdependence of observations as the null hypothesis unless proven otherwise. Gaussian methods and assumptions should only be used if the null is rejected.

CONCLUSION

Andriani and McKelvey (2000:16) say that "no statistical finding should be accepted into organization science if it gains significance via some assumption-device by which extreme events and (nearly) infinite variance are ignored," and O'Boyle and Aguinis (2012) assert that all existing theories of individual and organizational performance that have been tested using Gaussian techniques must be revisited. Both of these assertions, in my view, are too restrictive. However, scholars who ignore or disobey power laws in their empirical and theoretical studies of organization design do so at the peril of invalidity. Whereas it is of primary importance to account for power laws (and their generative mechanisms) while building and testing theories of organization design, the most rigorous, robust, and practically relevant theories will be crafted abductively. Here, stylized empirical facts from multiple sources – inductive field studies and deductive data analyses – can be integrated with computational simulation models of the generative process to develop explanations not only of what is but of what might be. This will provide scholars with the ability to develop prescient theory that can both explain the past and foretell the future (Corley & Gioia, 2011).

REFERENCES

Andriani P, McKelvey B. 2009. From Gaussian to Paretian thinking: causes and implications of power laws in organizations. *Organization Science* 20: 1053-1071.

Boisot M, McKelvey B. 2010. Integrating modernist and postmodernist perspectives on organizations: a complexity science bridge. *Academy of Management Review* 35(3): 415-433.

Carroll GR, Harrison JR. 1998. Organizational demography and culture: insights from a formal model and simulation. *Administrative Science Quarterly* 43(3): 637-667.

Corley KG, Gioia DA. 2011. Building theory about theory building: what constitutes a theoretical contribution? Academy of Management Review 36(1): 12-32.

Crawford GC. 2012. Emerging scalability and extreme outcomes in new ventures: power-law analyses of three studies. In L. A. Toombs (Ed.), *Proceedings of the Seventy-Second Annual Meeting of the Academy of Management*: ISSN 1543-8643.

Greene WH. 2007. *Econometric Analysis*. 6th ed. Prentice-Hall, Englewood Cliffs, NJ.

O'Boyle EH, Aguinis H. 2012. The best and the rest: revisiting the norm of normality of individual performance. *Personnel Psychology* 65: 79-119.

Popper K. 1963. *Conjectures and Refutations*. Routledge and Keagan Paul, London, UK.

Schoonhoven CB. 1981. Problems with contingency theory: testing assumptions hidden within the language of contingency "theory". *Administrative Science Quarterly* 26(3): 349-377.

Simon HA. 1955. On a class of skew distribution functions. *Biometrika* 42(3/4): 425-440.

Simon HA. 1968. On judging the plausibility of theories. In B. Van Rootselaar and J.F. Staal (Eds.), *Logic, Methodology, and Philosophy of Science III*. North-Holland, Amsterdam.

Starbuck WH, Nystrom PC. 1981. Why the world needs organizational design. *Journal of General Management* 6(3): 3-16.

West BJ, Deering B. 1995. *The Lure of Modern Science: Fractal Thinking*. World Scientific, Singapore.

G. CHRISTOPHER CRAWFORD

Ph.D. Candidate
Department of Management and Entrepreneurship
University of Louisville
E-mail: christopher.crawford@louisville.edu