# WORKING PAPERS

EAST-WEST CENTER WORKING PAPERS

EAST-WEST CENTER

The U.S. Congress established the East-West Center in 1960 to foster mutual understanding and cooperation among the governments and peoples of the Asia Pacific region including the United States. Funding for the Center comes from the U.S. government with additional support provided by private agencies, individuals, corporations, and Asian and Pacific governments.

*East-West Center Working Papers* are circulated for comment and to inform interested colleagues about work in progress at the Center.

For more information about the Center or to order publications, contact:

Publication Sales Office
East-West Center
1601 East-West Road
Honolulu, Hawaii 96848-1601

Telephone: 808-944-7145
Facsimile: 808-944-7376
Email: ewcbooks@EastWestCenter.org
Website: www.EastWestCenter.org

EAST-WEST CENTER WORKING PAPERS

EAST-WEST CENTER

Economics Series

No. 71, September 2004

# Limits to Modularity: A Review of the Literature and Evidence from Chip Design

**Dieter Ernst**

**Dieter Ernst is a Senior Fellow and Theme Leader** for Economic Studies at the East-West Center. His previous affiliations include the OECD, Paris, as senior advisor, and the Berkeley Roundtable on the International Economy (BRIE), at the University of California at Berkeley as research director. Dr. Ernst co-chairs an advisory committee for the U.S. Social Science Research Council (SSRC) to develop a new program on Asia. He has served on the Committee on Information Technology and International Cooperation (ITIC) of the SSRC. He also serves as advisor to the Committee established by the Prime Minister's Science Advisor to set up a National Science Foundation for Malaysia. Recent books include *International Production Networks in Asia: Rivalry or Riches?*, and *Technological Capabilities and Export Success: Lessons from East Asia*.

This paper has been prepared as part of the East-West Center's research project on Globalization of Knowledge Work: Why is Chip Design Moving to Asia.

# LIMITS TO MODULARITY
## - A Review of the Literature and Evidence from Chip Design

**by**
**Dieter Ernst,**
**East West Center, Honolulu, Hawaii, USA**
**ernstd@eastwestcenter.org**

**Draft2: 17 September, 2004**
**Word count: 12,224**

## Introduction

Debates on how business organization has moved beyond Chandler's vertically integrated multi-divisional firm (Chandler, 1977) have greatly benefited from the concept of "modularity" (e.g., Baldwin and Clark, 2000). We know that progress in the division of labor in design (technical modularity) has created opportunities for vertical specialization in project execution, enabling firms to disintegrate the value chain as well as to disperse it geographically. Increasingly this process has taken on a global dimension, giving rise first to offshore assembly in a wide range of industries (both traditional ones like garments and innovative ones like semiconductors), and later on to outsourcing to global contract manufacturers[1]. More recently, modular design has also provided ample opportunities for vertical specialization in the production of knowledge-intensive services, such as software, information services, engineering and R&D[2].

However, as so often happens in the history of social sciences, the success of an idea may sometimes lead to exaggerated claims that forget that even the best theories have limits. There is a tendency in the "modularity" literature to generalize empirical observations that are context-specific and to confound them with prescription as well as prediction. In this view, modularity becomes an over-arching business model that changes the way firms do business and how they organize to develop their products, processes and organization. Some proponents even claim to have discovered "a new American model of industrial organization" (Sturgeon, 2002). It is argued that modularity gives rise to a set of market-supporting institutions, notably stable interface standards and design rules, enabling firms to pursue focused strategies that rely heavily on outsourcing across the value chain. A resurgence of market forces that modularity facilitates will thus make the "Visible Hand" of large corporations invisible (Langlois, 2003). In short, modularity requires less coordination through corporate management.

In response to such strong claims of pervasive modularity, an alternative set of theories has emerged that center on the dynamics of modularity (Chesbrough, 2003b), "system integration" (Pavitt, 2003 a and b; Prencipe, Davies and Hobday, 2003), "open innovation" (Chesbrough, 2003a), and "iterated co-design" (Sabel and Zeitlin, 2004).

---

[1] On garments, see Gereffi (1994); on semiconductors, see  Ernst (1983); and on outsourcing, see Ernst and O'Connor (1992), Hobday (1995), Luethje (2002), Sturgeon (2002), and Ernst (2004a).

[2] For instance,  Arora, Gambardella and Rullani, 1998; Ernst, 2002a and 2004b.

This revisionist literature argues that current theories of modularity "predict too much" (Chesbrough 2003b: 193), and that, by confounding technical and market modularity, these theories exaggerate the empirical significance of the latter (Sabel and Zeitlin, 2004). As modularity theories pay insufficient attention to the dynamics of technical advances, they fail to explore the complex forces and constraints that shape processes of modularization, such as the increase of complexity with each transition to a new architecture, and the critical role of systems integration. Most importantly, if codification does not reduce complexity (which it fails to do when technologies keep changing fast and unpredictably), then the division of inter-firm collaboration will remain constrained, and requires more (not less) coordination through corporate management (Brusoni, Prencipe and Pavitt, 2001; Pavitt, 2003a and 2003b; Brusoni, 2003; Ernst, 2004b; and Tokumaru, 2004). In short, there are fundamental limits to modularity.

In this paper, I assess what we know about the limits to modularity and their impact on firm organization and industry structure. I focus on evidence from chip design, drawing on interviews during 2002 and 2003 with a sample of 60 companies and 15 research institutions that are involved in chip design in the US, Taiwan, Korea, China and Malaysia[3]. I first summarize "stylized" propositions of the modularity literature that are well-established, as well as predictions that are controversial. In part 2, I review important findings of the revisionist literature. In the rest of the paper, I describe three fundamental limits to modularity and relevant management responses. Part 3 highlights important differences in the coordination requirements of "project execution" (to design and produce an artefact, e.g. a chip) and of "technology development" (to produce the underlying knowledge bases). Part 4 documents constraints to interface standardization in the semiconductor industry. Part 5 examines the slow-down of innovation that may result from too strict an adherence to market modularity. Finally, part 6 examines the difficulties that management faces when it attempts to respond to the limits to modularity in chip design.

## 1. The Modularity Literature - Propositions and Predictions

The modularity literature has made an important contribution to the study of technical change and economic institutions. It demonstrates that progress in the division of labor in design (technical modularity) has created new opportunities for the organization of firms beyond vertical integration, and that this may transform industry structure and markets. The starting-point is technical change - "(m)odularity is a particular design structure, in which parameters and tasks are interdependent within units (modules) and independent across them." (Baldwin and Clark, 2000: 88). But the main concern is with implications for business organization and markets. Modularity thus is defined broadly to encompass

---

[3] The sample includes ten *strategic groups* of firms (both global and regional players) that participate in global design networks (GDNs): system companies; integrated device manufacturers (IDMs); providers of electronic manufacturing services (EMSs) and design services (the so-called ODMs, or "original-design-manufacturers); "fabless" chip design houses; "chipless" licensors of "silicon intellectual properties" (SIPs); chip contract manufacturers ("foundries"); vendors of electronic design automation (EDA) tools; chip packaging and testing companies; and design implementation service providers; and institutes and universities (both private and public). In China, the sample includes state-owned enterprises (SOEs), collective enterprises, and private technology firms.

all "techniques for dividing effort and knowledge …(that)… are fundamental to the creation of highly complex manmade things." (ibid: 5).

What precisely does this literature tell us about the link between technical modularity and the organization of firms and markets? It is of course impossible to do justice to the many extremely insightful findings of a still relatively young and thriving body of literature. For our purposes, I suggest to focus on three 'stylized' propositions that are well established through both theoretical and empirical work, and on two predictions that are controversial.

### 1.2. Propositions

The first proposition focuses on the convergence of technical and organizational modularity. It is argued that the architecture of a complex artefact corresponds with the organizational structure of the firm producing that artefact. The computer industry is frequently cited as an important breeding ground for this new industrial organization model (e.g., Langlois, 1992; Baldwin and Clark, 2000). Until the early 1980s, IBM personified 'vertical integration' within a multi-divisional firm: almost all ingredients necessary to design, produce and commercialize computers remained internal to the firm. This was true for semiconductors, hardware, operating systems, application software, and sales and distribution. Since then, modular design, based on standard interchangeable components as well as the widely shared Wintel architecture, has rapidly eroded the economic rational for vertical integration. This, it is argued, has made vertical specialization ("organizational modularity") the industry's defining characteristic.

A second proposition argues that organizational modularity is made possible by a combination of two developments: the codification of knowledge, enabled by the rapid diffusion of cheaper and more powerful information and communication technologies (ICT) (Sichel, 1997; Flamm, 1999); and market-led standardization (through technical standards and design rules) of the interfaces between organizationally separate stages of production. From a machine to automate transaction processing, the focus of ICT-enabled information management has shifted to the extraction of value from information resources, and then further to the establishment of Internet-enabled flexible information infrastructures that support not only information exchange among dispersed network nodes (e.g., Arora, Gambardella and Rullani, 1997), but also the sharing, utilization, and creation of knowledge among multiple network participants at remote locations (Jørgensen and Kogstie, 2000; Foray and Steinmueller, 2001).

A third proposition highlights implications for competitive strategies. It is argued that codification, together with shared interface standards and design rules reduces the "volume of information", and hence the amount of knowledge sharing, that is required for inter-firm coordination. This makes it possible to transform products into fully "modular" or decomposable building-blocks, enabling firms to pursue focused strategies that rely heavily on outsourcing across the value chain. A computer company can now focus on those activities ('core competencies') that generate the highest margins and which are critical for sustaining the company's competitive advantage. In turn, the company outsources those low-margin activities in which it does not have a strategic advantage (e.g., Ernst, 2002b). While outsourcing initially was focused on manufacturing, it now extends to all stages of the value chain, including research and new product development (Chesbrough, 2003a; Ernst, 2003 and 2004b).

In short, the afore-mentioned three propositions of the modularity literature explore how progress in design methodology ("technical modularity") has created new opportunities for vertical specialization ("organizational modularity) in project execution, enabling firms to disintegrate the value chain as well as to disperse it across firm boundaries and geographic borders. These propositions have been corroborated by the experience since the 1980s of a rapid expansion of global outsourcing.

### 1.1.Predictions

However, the modularity literature moves into murky waters, when it attempts to generalize empirical observations that are context-specific, and when it ventures to provide prediction as well as prescription about the evolution of technology and industry structure. For instance, much of the current modularity literature shares the assumption (at least implicitly) that every technology will proceed from a less modular, more integrated state towards a more modular state (e.g., Baldwin and Clark, 2000). In this view, modularity becomes the stable end state of industry evolution, and this is claimed to be true across industries and technologies.

A related equally strong prediction addresses the evolution of industry structure. It is argued that technical modularity leads to a resurgence of market forces, and that this will make the "Visible Hand" of large corporations invisible (Langlois, 2003). In this view, technical modularity is expected to lead to a new type of division of labor - a proliferation of modules produced by specialized firms using general-purpose tools (both machinery and software). As the market is the primary driver of this decentralization of economic transactions, less coordination is required through corporate management. At the same time, this will also erode entry barriers, and hence will undermine existing oligopolistic market structures. Two role models are mentioned in the literature: US-based global providers of electronic manufacturing services (EMS) which for Sturgeon (2002) represent a "new American model of industrial organization"; and Silicon Valley-style innovative start-up companies that focus on high-margin market segments, attract ample venture capital funding, and hence grow rapidly beyond their initial small size (e.g., Arora, Gambardella and Rullani, 1997).

## 2. The Revisionist Literature

However, empirical research on a variety of industries (both highly innovative and traditional ones)[4] has raised important questions to what degree the afore-mentioned predictions can be sustained. This has given rise to an alternative set of theories that center on the dynamics of modularity (Chesbrough, 2003b), "system integration" (Pavitt, 2003a, and Davies, 2003), "open innovation" (Chesbrough, 2003a), and "iterated co-design" (Sabel and Zeitlin, 2004). This revisionist literature highlights three specific weaknesses of current theories of modularity: they confound technical with market modularity; they exaggerate the empirical significance of market modularity; and they fail to explore the dynamics of modularization. These weaknesses need to be addressed to capture the complex forces and constraints that link technical modularity with changes in organization and market structure.

---

[4] Studies of highly innovative sectors include aero-engines (Prencipe, 1997), telecommunications infrastructure (Davies, 1997), hard disk drives (Chesbrough, 2003b), and chip design (Tokumaru, 2004 and Ernst, 2004b). Studies of traditional sectors cover chemical engineering (Brusoni, 2003), oil exploration (Acha, 2002), automotives (Takeishi and Fujimoto, 2003) and tyres (Acha and Brusoni, 2002).

### 2.1. Technical versus market modularity

Chesbrough's distinction between technical and market modularity (2003b) highlights an important weakness of the modularity literature: a failure to distinguish between technical modularity and related changes in institutions, i.e. firm organization and markets. "Market modularity" is used as a shorthand for the institutional, informational and cognitive conditions that "enable markets to take over…( at least parts of, DE) … the coordination tasks of innovation." (ibid. : 178) Both technical and market modularity may go hand in hand in some cases, but they need not do so. Technical modularity is a necessary, but not sufficient condition for market modularity. Chesbrough highlights four criteria that must be met for both types of modularity to converge: knowledge diffusion, shared language, verification, and supplier base.

Market modularity implies that, instead of a vertically integrated firm (e.g., IBM), multiple specialized suppliers contribute components (both hardware and software) that constitute a specific architecture. All of these firms thus need to be able to share the knowledge of how the components interact within a particular design architecture[5]. This implies that the knowledge of how to implement technical modularity through adjustments in organization and management procedures, must now be diffused outside the firm (e.g., through open standards and production networks). Second, for network participants to be able to clearly communicate their requirements, there must be a shared language that can be used to specify unambiguously the features and functions of components. In other words, network participants must agree on a standard language that enables them to share their diverse concepts and codes.

Third, of particular importance is the verification criterion: tools and equipment must exist to verify that the required attributes of components have been met. In chip design, for instance verification has become a critical bottleneck that constrains market modularity[6]. Finally, while technical modularity facilitates the development of complex networks for manufacturing and design, there is no guarantee that technical and market modularity will converge. For this to happen, a capable and broad-based supplier base must exist that allows the "flagship" of the resultant networks (Ernst, 2002b) to switch suppliers.

Using these four criteria can help to determine for a particular industry whether modular design ("technical modularity") has been transformed into market modularity. Take the semiconductor industry where shared design rules have initially helped to separate design from fabrication, juxtaposing fabless design houses and silicon foundries in a modular market structure. However, the simultaneous transition to submicron

---

[5] Computer designers use the term "architecture" to refer to "the partitioning of the … (computer) … system into components of a given scope and related to each other functionally and physically through given interfaces. From a given architecture flows the design of components' functions and how they relate to each other…" (Gawer and Cusumano, 2002: 18). These authors compare a system architecture with a geographic map, where the components of the system are the countries' territories and the interfaces between components are the countries' borders (ibid.:19).

[6] With growing design complexity, it becomes necessary to verify early and frequently whether the SoC design can be produced at sufficiently high yield, and whether it will do what it is expected to do. Today, 60 to 70% of SoC hardware design time goes into verification, leaving only 30 to 40% for the actual device development. This "verification crisis" obviously constrains considerably the productivity of design. " (Ernst, 2004b: p.11 WP)

process technologies and system-level integration design (SoC and SiP) has dramatically increased complexity, with the result that the first three of the above four criteria for market modularity are no longer met (Ernst, 2004b): intractable constraints to interface standardization obstruct knowledge diffusion and the use of a shared language to specify unambiguously design rules and the features and functions of components. At the same time, the "verification bottleneck" has been spiraling out of control.

In short, we need empirical research to identify precisely "…when and why technical modularity leads (or does not lead) to modularity in the market." (Chesbrough 2003b: 194).

### 2.2. Limited empirical significance of market modularity

On the empirical significance, the electronics industry is often taken as the pre-eminent example of market modularity. Yet, as Sabel and Zeitlin (2004: 10) rightly emphasize: "arm's-length coordination among specialists based on standard technical interfaces… (i.e. "market modularity", DE) is less common than is often claimed"[7]. For instance, in 2002, the share of U.S. based contract manufacturers (the providers of electronic manufacturing services, or EMS) in worldwide production of electronics hardware was estimated to be around 16%[8].

But even more important are fundamental changes in the organization of electronic contract manufacturing. Ernst (2004c) highlights serious limitations to the US model of contract manufacturing. For instance, global contract manufacturing is a highly volatile industry. While powerful forces push for outsourcing, this process is by no means irreversible. Major global brand leaders (the so-called original-equipment manufacturers, OEMs) retain substantial internal manufacturing operations; they are continuously evaluating the merits of manufacturing products or providing services internally versus the advantages of outsourcing (e.g., Bengtsson and Berggren, 2002). Second, global EMS are now in a much weaker bargaining position than OEMs, whose number has been reduced by the current downturn and who are now much more demanding (e.g., Benson-Armer et al, 2004).

Furthermore, serious conflicts of interest have emerged between OEMs, who are looking for flexibility, and EMS, who are looking for predictability and scale. For instance, OEMs focus on early market penetration and rapid growth of market share to sustain comfortable margins. OEMs thus need flexibility in outsourcing arrangements that allows them to divert resources at short notice to a given product as it becomes a hit. This sharply contrasts with the situation of EMS: with razor-thin margins, they need to focus ruthlessly on cost cutting. EMS need predictability: "they want to make commitments in advance to reap benefits like big-lot purchases and decreased overtime." (Lakenan et al , 2001, p.10).

These conflicting interests complicate the coordination of these "modular" production networks. They also require substantial changes in the organization of both OEMs and EMS, as well as an alignment of incentives through contract terms and agreements. The

---

[7] The limited empirical significance of market modularity has also been documented by research on complex mechanical assembly industries (such as automobiles, aircraft , construction and agricultural machinery) and chemical engineering (see sources quoted in Sabel and Zeitlin, 2004 and in Prencipe et al, 2004).

[8] Ernst, 2004a, quoting data provided by Eric Miscoll, CEO, Technology Forecasters, Inc, April 15, 2002.

irony is that, the more "modular" these contract manufacturing networks become, the more difficult it is to coordinate them.

In addition, the proponents of modularity have neglected a powerful trend towards reintegration of manufacturing and design that appears to counter earlier modularization processes. Let us first look at manufacturing. Since the turn of the century, two different business model compete for success in the global electronics contract manufacturing industry - EMS (= electronic manufacturing services) and ODM (= original design manufacturing)[9]. The two main strengths of EMS companies are process-related, with a focus on scale and global presence and "integrated solutions". But EMS are typically weak in product-specific knowledge and design capabilities, which keeps their profit margins low. Contrast this with the ODM business model which provides reasonable margins, but requires good design capabilities. More specifically, ODM companies must be able to choose the appropriate hardware platforms, and they must be able to integrate software and applications of growing complexity, while meeting a diverse array of standard protocols. Equally important are short design cycles and a capacity for flexible response. This must be combined with sufficient scale, sophisticated component sourcing and a strong supply chain in low-cost Asia.

Both EMS and ODM have tried to overcome their respective disadvantages, giving rise to an at least partial convergence of EMS and ODM business models. No EMS can survive without covering ODM services, and vice versa. Economies of scale and scope determine success in this business, which drives the growth of these firms, primarily through mergers and acquisitions. As EMS grow and expand globally, system integration capabilities become their most important differentiating factor.

In short, it is fair to argue that "market modularity", defined as reliance on arms'-length transactions mediated by market forces, has provided new opportunities for global brand leaders to pursue selective strategies of vertical specialization. It would however be misleading to claim that "market modularity" dominates electronics manufacturing.

That claim can also no longer be made for chip design. During the initial euphoria of the late 1990s, there were widespread expectations that technical modularity, i.e. the reuse of design knowledge through modular design, combined with the transition to system-on-chip (SoC) design, would naturally lead to organizational and market modularity. This perception is mirrored in the claim by Linden and Somaya (1993: 545) that "we observe a burgeoning market for licensed DMs (= design modules, or SIPs in industry parlance, DE)". In the conclusions of this article, the authors are struggling to support this claim. Linden and Somaya (2003: 571) state that "integrated modes currently enjoy the upper hand in SoC, primarily due to the initial distribution of industry assets and the lack of supporting licensing institutions at the dawn of the SoC era." They argue however that the balance may shift over time towards open market SIP licensing.

Yet, this prediction is not supported by empirical evidence. There are three external sources of SIPs for SoC design teams: (1) in-house SIP libraries that compile building blocks from previous designs; (2) other in-house design teams working in parallel with

---

[9] The term "EMS" is used for companies that provide manufacturing (PCB assembly and box assembly) and order fulfillment services, across a wide range of product markets, to global system/brand name companies (the so-called OEM customers). The term "ODM" on the other hand describes companies that provide product design, manufacturing and distribution logistics for OEM customers with a relatively narrow product focus.

the main SOC design team; and (3) independent SIP vendors. Of these, independent SIP vendors are the least important source. Instead, the exchange of SIPs takes place primarily *within* global system companies and flagship-dominated global design networks (GDNs) (e.g., Martin, 2003:11; Goering, 2002). The global market for SIPs has remained relatively small (Ernst, 2004b, figure 6, The Global Market for SIPs). It consists primarily of a few "star" SIPs (for MPUs, DSPs, memories and analog design blocks), dominated by a handful of specialized global suppliers like ARM, MIPS, Rambus, and DSP Group.

In short, market modularity has not come to dominate chip design. In contrast to predictions of a convergence of technical and market modularity in that sector, organizational integration appears to gain in importance, albeit in novel hybrid forms that differ from Chandler's vertically integrated firm. Ernst and Martin (2004) document that, propelled by increasing complexity in process and design technology, a shift is under way in market power to those system companies and integrated device manufacturers (IDMs) that own system knowledge, SIPs, as well as advanced process development skills. On the other hand, favorites of the "modularity" literature, such as foundries, fab-less design houses, chip-less SIP suppliers and EDA tool providers are facing new and very demanding challenges to their established business models.

All of this indicates that the maturity of a technology may be a necessary condition for the convergence of technical and market modularity. This brings us to the dynamic aspect of the modularity story.

### 2.3. Dynamics of modularization

Current modularity theories pay insufficient attention to the dynamics of technical advances, and hence fail to explore the complex forces and constraints that shape processes of modularization. Of particular importance are cognitive limits to modularity that emerge with each transition from a well-established ("mature") to a new ("fluid") design architecture. Proponents of modularity have left unexplored what happens when component and architectural design, as well as fabrication technologies keep changing fast and unpredictably, and how that affects the convergence of technical and market modularity.

Take again chip design (Ernst, 2004b). Since the mid-1990s, intensifying pressures to improve design productivity, combined with increasingly demanding performance requirements for electronic systems (lighter, thinner, shorter, smaller, faster, multi-functional, less power-consuming, and cheaper) have produced an upheaval in chip design methodology[10]. Under the label of "system-on-chip" (SoC) design, systematic attempts have been made to enhance design methodology by combining "modular design" and "design automation" with attempts to move design from the individual component on a printed circuit board closer to "system-level integration" on a chip (Chang et al, 1999 and Martin. and H. Chang, eds., 2003). Fed by a massive inflow of abundant venture capital during the "New Economy" boom, which has encouraged

---

[10] "Design methodology" is the sequence of steps by which a design process will reliably produce a design "as close as possible" to the design target, while maintaining feasibility with respect to constraints. All known design methodologies combine (1) the enforcement of system specification and constraints via top-down planning and search, with (2) a bottom-up propagation of constraints that stem from physical laws, limits of design and manufacturing technology, and system cost limits.

aggressive fabless semiconductor startups, this has loosened the bonds between design and fabrication.

Best known is the symbiotic relationship between the two main carriers of modularity: "fabless" design houses that outsource chip fabrication, and specialized suppliers of contract fabrication services, the so-called "silicon foundries". Less well known, but equally important are a diverse group of specialized providers of complementary services, such as tools for electronic design automation (EDA), verification, testing, and the development of embedded software. It is this de-coupling of design and fabrication and its fragmentation into diverse specialized suppliers that gave rise to the afore-mentioned predictions of a convergence of technical and market modularity.

Note however that Macher, Mowery and Simcoe (2002) emphasize the persistent tension between vertical specialization and re-integration. This argument is supported by more recent research. In fact, current transformations appear to move away from market modularity in the opposite direction to more integrated, albeit new hybrid forms of industrial organization (Ernst and Martin, 2004). The earlier de-coupling of design and manufacturing is now being reversed by the increase in cognitive complexity (both on the chip/"silicon" and on the "system") that results from advanced process geometry and progress in SoC design. This constrains the division of inter-firm collaboration within global design networks, and requires more (not less) coordination through corporate management. Large global corporations (the network flagships) are required as "knowledge integrators" to coordinate multiple interfaces between specialized bodies of knowledge within global design networks that have evolved as a result of technical modularity.

Chesbrough's dynamic theory of modularity provides a useful framework for analyzing how each transition to a new architecture affects the process of modularization (Chesbrough,. 2003b; Chesbrough and Kusunoki, 2001). Chesbrough argues that the evolution of technology is cyclical, moving from "interdependent" to "modular" design architectures and back to "interdependent" ones, and that this technology cycle conditions the choice of organization and strategy. This reflects fundamental "limits to modularity" that are most prominent at the early stages of the development of a new technology. It is worthwhile to quote in detail this proposition: "When technologies are interdependent, internal organization can better sort out these complex technical interactions than market-mediated transacting, due to the lower coordination costs of internal organization….This internal advantage goes away when technologies' interrelationships become better understood…. (T)he incentives available within the market… (now) elicit greater innovation, and spread the costs of innovation across a wider market, relative to internal organization. At some later time, though, once these well-understood, modular technologies hit their performance limits, an interdependent technology phase arises anew." (Chesbrough 2003b: 184-85).

Chesbrough's model distinguishes three stages: interdependent architecture 1 (I1), modular architecture 1(M1), and interdependent architecture 2 (I2). Technological interdependence characterizes any new technology. During stage I1, "the requisite information of how the different elements function together is not well defined, and interactions between elements are poorly understood….. In this early stage of technology evolution, managerial coordination, rather than markets, provides the most effective mechanisms to coordinate the relationship between elements of the system." (Chesbrough, 2003b: 176).

With growing maturity of the technology, modularization becomes possible, leading to stage M1. Engineers can now use the twin tools of analysis and synthesis to reduce significantly the complexity of the technology interfaces (Simon, 1962; von Hippel 1990; Kogut and Zander, 1993; Henderson and Clark, 1990). When benefits from established modular architecture are decreasing, the search for a new design architecture will lead to a new stage I2. At each stage, firms pursue different types of learning: learning-by-doing is required during I1, while learning-before-doing (Pisano, 1996) becomes possible during M1. The transition to I2 requires re-learning and a return to learning-by-doing.

Chesbrough (2003b) uses the development of heads, one of the key components of a hard disk drive, to demonstrate the dynamic cyclical evolution of modularity. He shows that, with each transition to a new generation of head technology, integrated modes of organization have regained advantages over more modular modes of organization - at least for a certain period of time. With the transition from iron-oxide to thin-film heads, "…(t)he integrated head manufacturers (i.e. the firms that made drives, as well as heads) were able to adopt the thin film heads years ahead of the merchant suppliers, precisely because they could employ internal administrative means to coordinate the interdependencies between the heads and the drive design." (Chesbrough 2003b: 183).

A similar shift occurred with the development of a recording head technology, called *magneto-resistive* (MR) heads. Extreme interdependence is the hallmark of this complex technology - "the design of the disks, actuator mechanisms, and read-write channels depended upon the design of the head - and vice versa." (Chesbrough 2003b: 183). Again, integrated manufacturers who made their own heads such as IBM, Hitachi and Fujitsu, found it easier to sort out these interdependencies, and figure out solutions to these problems. Chesbrough and Kusunoki (2001) demonstrate that non-integrated disk drive companies like Western Digital and Maxtor in the US, and NEC and Toshiba in Japan, found themselves in a "modularity trap" as they lacked the internal systems integration capabilities to cope with the interdependencies of MR heads.

This pattern also holds for the example of chip design that I have introduced at the beginning of this section. As both design and fabrication technologies keep changing fast and unpredictably, the resultant increase in the cognitive and organizational complexity of chip design has produced an extreme degree of interdependence. Consequently, for each SoC design project, global design networks are constructed around project leaders (the network flagships) with systems integration capabilities strong enough to coordinate multiple design interfaces. These knowledge integrators are needed to coordinate multiple interfaces between fabless design houses and foundries, as well as with a diverse group of specialized providers of complementary services, such as tools for electronic design automation (EDA), verification, testing, and the development of embedded software (ESW).

Of course, pure modularity and pure technological interdependence are extreme boundary conditions (Brusoni and Prencipe, 2001) - most products and technologies exist somewhere along a continuum between these extremes. But these concepts are useful for examining the arduous process of advancing to a new, and better architecture in an industry characterized by significant modularity. Chesbrough's dynamic theory of modularity demonstrates that, if a firm fails to adjust its organization and innovation management to the requirements of the new architecture, it risks being caught in a "modularity trap." In other words, if a firm focuses too much on developing products *within* given interface standards, this may erode the firm's system integration capabilities. A "modularity trap" exists, when flagships fail to retain those system integration capabilities that are necessary to incorporate new (interdependent) component technologies effectively into their systems (Chesbrough and Kusunoki, 2001).

In short, a dynamic perspective of modularization processes shows that almost every technology gives rise to a "modularity trap" at some stage in its evolution. This is so because attempts to advance beyond the current architecture reintroduce "the complexity at the systems level that modularity was intended to manage." (Chesbrough 2003b:190). It is thus difficult to sustain the assumption, implicit in much of the modularity literature, that modularity is the stable end state of industry evolution, and that this is true across industries and technologies (Baldwin and Clark, 2000). While modular design has acted as a powerful catalyst for changes in business organization and industry structure, limits to modularity are aplenty, and constrain the convergence of technical, organizational and market modularity.

## 3. Limits to Modularity (1) - Demanding Coordination Requirements

That this is so should not be surprising. As Pavitt (1999:XX) has convincingly argued, activities that require complex knowledge pose very demanding coordination requirements. There are cognitive limits to the process of modularization. Based on Brusoni (2003) and Tokumaru (2003), it is possible to highlight important differences in the coordination requirements of "project execution" (to design and produce an artefact, e.g. a chip) and of "technology development" (to produce the underlying knowledge bases).

Baldwin and Clark (2000: chapter 3) correctly emphasize that modularity in design has created opportunities for vertical specialization (combining disintegration and geographic dispersion) in project execution. Their analysis however neglects the increased knowledge exchange that is necessary to develop design and manufacturing technologies. This, in turn, requires ex ante coordination through integration in technology development. Modular product design thus needs knowledge-integrating firms to coordinate specialized bodies of knowledge and increasingly distributed learning processes. It does not reduce the need for system integration.

This argument runs counter to established wisdom. Arora, Gambardella and Rullani (1998) for instance argue that modularity can be applied to innovation as well as to manufacturing. "This is because modularity allows for a separation of the innovation process in two main activities: The production of basic (standardised) modules, and their combination to produce variants of technologies or product designs that are better suited to the special needs of individual users or markets." (ibid: 123). In this view, there is a clear-cut separation between specialized suppliers and system integrators, with interactions between these firms coordinated by the market. Each supplier, focusing on a specific innovation module, would be able to specialize in specific learning and innovative efforts, while the "combination of modules" will be performed by downstream firms or by the users themselves. This "modular" division of labor is made possible by the use of information technology (Arora, Gambardella and Rullani mention simulation software for design purposes) that allows to codify knowledge, and hence facilitates knowledge exchange. Modularity is supposed to erode the importance of ownership control, because it creates alternative market-based coordination mechanisms. This is so because a modular architecture creates an ‘information structure’ that cannot be changed during the duration of the project, and hence can be used to coordinate decentralized design teams.

But this argument fails to acknowledge that the reuse of knowledge to combine modules requires a broader knowledge base for systems integration. In other words, modular product design may well increase complexity and hence the need for system integration. Large global network flagships retain diversified technology bases precisely to cope with the demanding coordination requirements of disintegrated and geographically dispersed “technology development”. This is in line with the findings of Granstrand, Patel, and Pavitt (1997) that large firms are more diversified in the technologies they master than the products that they make and that their technological diversity has been increasing while typically their product range has narrowed.

There is no automatic link between adopting a modular product design strategy and the adoption of a modular organization. Such a direct link is even less likely between the use of a modular product design and the modularization of the firm’s knowledge base (Brusoni, 2003: 5). In fact, it is necessary to push the argument one step further (Ernst, 2004b). Network flagships that can serve as “system integrators” are required to coordinate multi-layered global networks of firms that interact to conceive, design and engineer an artefact. As explained above, this is especially the case for new, “interdependent” design architectures - their development requires flagships that possess the systems integration knowledge to evaluate how best to advance across the myriad possible combinations (Brusoni, 2003; Brusoni and Prencipe, 2001).

Take again chip design (Ernst and Martin, 2004). By combining advanced process development, design IP and applications knowledge, global network flagships have been able to compete successfully against the “modularity” model, as represented by the collaboration between fabless design houses and foundries. This is certainly the case for integrated device manufacturers (IDMs) like Intel, Texas Instruments, and STMicroelectronics that have managed to develop “platform leadership” strategies. A combination of process technology, design IP and system applications knowledge have also helped system companies like IBM, Philips, Nokia, Samsung, LG and the leading Japanese electronics system companies. Even when they purchase semiconductors from specialized suppliers (whether from IDMs, like Intel or TI, or from fabless and chipless companies), these system companies make a strong effort to retain considerable internal expertise in SC process technology, fabrication, EDA tools, and design IP. They realize that, without in-house mastery of such capabilities, it would be difficult to retain systems knowledge - the latter being a necessary prerequisite for turning these assets into differentiating features of their products and services.

Of critical importance for the flagships of global design networks is the ability to frame the problems (architectures; components; performance features) to be addressed in specific, temporary design projects[11]. As explained by Brusoni (2003), flagships need to be able to identify the crucial technological and organizational interdependencies, to

---

[11] A project is defined as “basically a contract, or rather a bundle of contracts, that specificy the duties, responsibilities and rewards of the firms involved..” (Brusoni, 2003: 15).

explore alternative ways of de-composing the problem, and to isolate those options that do not deserve to be pursued further. They need to be able to identify, probe and freeze key interdependencies. For each specific project, i.e. the design of an embedded microcontroller for a mobile phone for Motorola, the network "provides a temporary administrative framework within which some form of hierarchical coordination replaces the market." (Brusoni, 2003: 15).

In short, modular product design may lead to vertical specialization in project execution, but it requires integration in technology development. Ownership of tangible assets is not a sufficient condition for the effective coordination of multi-layered global design networks. Equally important are network flagships that can serve as "system integrators".

## 4. Limits to Modularity (2) - Constraints to Interface Standardization

A surprising feature of modular systems is their considerable rigidity. Once deployed, interface standards are difficult to adjust. When the architectural yield achieved from a given relationship of components approaches its limit, it becomes necessary to establish a new architecture. But a defining characteristic of modular systems is that any transition to a new generation of design architecture requires fundamental changes in system components, which consequently will break down established interface standards (Chesbrough, 2003b).

Take the afore-mentioned transformation of chip design, where a re-coupling of design and chip fabrication requires fundamental changes in design methodology. The result is that established interface standards cease to function, even for the most profoundly documented design rules that link fabricators (the silicon foundries) and design houses (Ernst, 2004b). Another example is the transition from 8-bit to 16-bit microprocessors that required massive changes in complementary software and hardware, and hence necessitated a redefinition of interface standards (Chesbrough, 2003b: 180, 181).

Sabel and Zeitlin (2004: 2m) demonstrate that it is impossible to establish "…standard design interfaces so comprehensive and stable that customers and suppliers can in effect interact as if operating in spot markets for complex components or subassemblies without jeopardizing their long-term survival." In line with Pavitt (1999), they argue that innovation creates "unpredictable novelty", requiring continuous relearning. Hence, "modularization, at least in the form of entrenched standards, cannot be." (Sabel and Zeitlin, 2004: 5 m).

Chip design provides an important example of the tight limits to interface standardization. Based on standard interfaces and design rules, the division of labor used to be reasonably simple during much of the 1990s. The resulting separation of chip design and fabrication has been one of the favorite example of modularization proponents. Engineers designed chips and handed the definition to the mask makers, who then sent the masks to the wafer manufacturers (the silicon foundries). And (most of the time, at least) the result of having this modular division of labor was a chip that could be manufactured at an acceptable yield.

But this easy phase of modularization of the semiconductor industry has vanished for good (Ernst, 2004b; Ernst and Martin, 2004). As process technology has dramatically increased in complexity, intense interactions are required across all stages of the semiconductor value chain, and it is no longer possible to work with entrenched standard interfaces and design rules. All participants in the semiconductor industry know that they need to find a way to organize collective and integrated solutions. They also know that uncertainty makes this extremely difficult, as does the fact that the industry is now vertically specialized.

Current attempts to develop a "Unified Data Model"(UDM) provide an excellent example of the limits to interface standardization in the semiconductor industry. UDM is a data base of chip definitions that encapsulates the information required for chip design, mask making and wafer manufacturing and that can function as a single data representation for a wide range of "electronic design automation" (EDA) tools. It is expected that UDM will enable designers to simulate a manufacturing run for different variations of a chip design and then address trade-offs between functionality, density and yield. Equally important benefits are expected to accrue to mask makers and wafer fabricators[12].

While potential benefits are substantial, two fundamental constraints are currently blocking the development of UDM: intense competition among oligopolists; and uncertainty that results from growing complexity. First, intense competition among oligopolists prevents companies from agreeing on a common standard. To understand why, let us first look at earlier attempts to develop common data formats for EDA tools. Conditions seemed to be favorable. There were strong pressures to move forward, as design engineers needed common data formats for EDA to collaborate more effectively. In addition, process and design technologies were relatively mature, enabling modular design architectures. Nevertheless, developing a common data base was not possible. Reflecting the oligopolistic market structure of the EDA industry, two competing data formats for EDA tools have emerged: "Open Access", led by Cadence, and supported by IBM, and "Milkyway", led by Synposis.

Developing UDM may be even more difficult. Fundamental implementation barriers indicate persistent limits to modularity. These limits result from a powerful combination of rising complexity due to a shift to a new, interdependent design architecture (SoC) and intense oligopolistic competition. As for the latter, current attempts to develop the UDM standard are based on the "Open Access" standard, which means that Cadence as well as IBM would greatly benefit if UDM would materialize. It is thus hardly surprising that both these companies are strong supporters of the UDM standard, while both Synposis and Mentor remain hostile[13].

But more important are complexity-related constraints that drastically increase uncertainty, obstructing the development of entrenched standards. Increasing complexity results from changes in design architecture as well as from advanced process geometry that pushes the limits of semiconductor fabrication. To cope with increasing complexity requires dense interactions between IC design and fabrication, and these interactions must occur simultaneously at multiple levels. In contrast to the earlier separation of

---

[12] Main carriers of UDM include the Silicon Integration Initiative Inc. (SI2), an organization of industry-leading semiconductor manufacturers, electronic system companies, design houses and EDA tool vendors, with the mandate to improve the productivity of integrated circuit manufacturing and design. This industry organization has started to incorporate work from the "Design to Mask Coalition"(DTMC) to extend the "Open Access" standard for EDA tools ( see below) with UDM-like features.

[13] Cadence actually provided the Open Access original code to the SI2 organization

design and manufacturing, current transformations appear to move in the opposite direction, to a more integrated model.

Design teams now must share data and exchange knowledge with mask makers and wafer fabricators (foundries). For instance, mask makers must have access to design data so that they can make quick changes in their techniques (such as for recticle enhancement), which gets increasingly tricky as component size shrinks. However, today's chip design methodologies do not provide mask manufacturers with the data they need. As a result, mask makers "…often waste time tuning circuits inappropriately or, worse, tuning layout features that aren't even part of the actual circuitry…. There's always the danger that the tweaking that takes place during mask making may result in accidental changes to the performance of chip circuits, necessitating expensive design iterations. " (James, 2004: 2).Design rules are also no longer capable of coordinating the interface between design and fabrication. With growing miniaturization, design rules become geometrically more complex and interdependent. The chief technical officer of a major photomask producer worries that "… at 65nm, designers may end up with so many rules that it will be impossible for a mere mortal (even a design engineer) to cope with all of them." (quoted in James, 2004: 2).

All of this indicates very tight constraints to interface standardization. "We simply don't have enough science to understand how to create a data format that would work in all these different environments." (Joe Sawicki, VP of the design-to-silicon group at Mentor, as quoted in James, G., 2004: 3). Other observers go even further, and argue that a common interface standard is logically impossible, as "… the tasks and tools in different (design) phases need to see data very differently to be effective…(Hence), (e)ach tool requires a data model that is specific to the task." (Richard Goldman, VP of strategic market development at Synopsis, as quoted in James, 2004: 3). An additional modularity constraint is built-in obsolescence, a concern voiced especially by IP providers. In this view, common standard interfaces are much too rigid to cope with frequent adjustments in chip manufacturing processes which are commonplace at the sub-micron level.

Our analysis of constraints to interface standardization in chip design confirms Chesbrough's dynamic theory of modularity (see 2.3. of this paper). These constraints result from the transition to new and immature architectures and process technologies. I part however company with Chesbrough's analysis, in emphasizing that these limits to modularity will not necessarily lead to a pendulum-like swing back to the *status quo ante* of vertical integration. Instead, as demonstrated in Ernst and Martin (2004), new hybrid forms of industrial organization that mix and match elements of modularity and integration, are likely to emerge. These hybrid organizational responses reflect the simple fact that firms need to cope with highly complex technical and competitive challenges for which no ready-made organizational solutions exist.

Take design rules. With increasing complexity, they need to treat a fine balance. Design rules need to be "enabling, i.e. firm enough to encourage modular innovation and recombination – but loose enough not to be constraining to the evolution of the system." (Sabel and Zeitlin, 2004: 5,6 m). This is why inter-firm collaboration increasingly relies on emerging methodologies and practices of iterated co-design, such as concurrent engineering, bench-marking, co-location of personnel, problem-solving teams, and processual quality standards. This forces "collaborators to question and clarify their assumptions about their joint project,… (allowing) for a corrigible partitioning of tasks within and across firm boundaries (, DE: and national borders) in ways that fixed modular interfaces do not." (Sabel and Zeitlin, 2004: 2 m).

The only thing certain is that, with rising complexity, designers, mask makers and wafer fabricators must move beyond the "modular" design network. They must find new and unconventional ways to work together "collaboratively" as part of more integrated global design networks. For this to happen, however, all three actors will need a shared data model, so redesigns that are mandated during fabrication will not force the designers to start again from scratch. But developing such a common data base for interdependent design architectures is extremely difficult.

Whether these limits to modularity can be overcome, depends in the end on competitive dynamics. There might be forces that can break the limits of oligopolistic rivalry. For instance, a leading foundry (e.g. TSMC) or integrated deveice manufacturer (IDM) could impose a solution by forcing the rest of the industry to use whatever UDM definition would work well in its factories. But this would force the foundry or IDM to reveal more about its own internal data representations than it might normally want to do. Alternatively, a "customer "revolt" by the chip design community may help to break the stalemate: "… at some point, EDA users will probably… refuse to buy additional tools until they get the interoperability that's necessary to preserve their cost structure." (James, 2004: 5).

## 5. Limits to Modularity (3) - A Slow-Down of Innovation

This paper has highlighted inherent rigidities of pure-play market modularity. There are thus strong reasons to belief that market modularity, without corrective moves towards organizational integration, may slow down the pace of innovation. First, as demonstrated by Chesbrough (2003b: 180), modular systems are hard to advance. "While component innovation within the system can continue to occur so long as the boundaries of the components' relation with the rest of the system are respected, innovation at the system level becomes increasingly problematic…. Thus, the very partitioning of the system that enabled market modularity to develop, later turns into a restriction upon that system's further evolution."

Second, advances in modularity may reduce the rents that firms can capture from innovation. In industries like hard disk drives (HDDs), PCs and consumer electronics, high modularity apparently has led to a loss of system integration capabilities. As for HDDs, during phase I1, "IBM profited not only from its use of a more advanced component, it profited as well from its ability to manage the interactions of the advanced component with the rest of the system." (Chesbrough, 2003b: 191). But once the industry moved to phase M1, IBM was no longer able to reap rents from defining the systems architecture (through closed proprietary standards)[14].

But there is also a third reason why modularity may slow down the pace of innovation. Even if flagships retain extensive systems integration capabilities and market power, they may find it difficult to advance architectures ("lead platforms") in an industry, precisely

---

[14] I would argue however that this is not the end of the story, and that one should not restrict the analysis to a single product market. Once IBM understood that its rent generation capability for HDD was declining, it began to diversify into other areas, especially IT services. This explains why IBM eventually sold its HDD operations to Hitachi.

because so much modularity has been developed. A typical example are the difficulties faced by Intel in its attempt to promote its "platform leadership" strategies (Gawer and Cusumano, 2002: 39-76). These strategies are defined by decisions on the "system architecture (the degree of modularity), interfaces (the degree of openness of the interfaces to the platform), and intellectual property (how much information about the platform and its interfaces to disclose to outside firms)" (ibid: 40). To build a coalition of complementors to its platform, Intel must shoulder the substantial costs and risks involved in making new products that complement the new architecture.

This gives rise to an important strategic dilemma that has been widely emphasized in the revisionist literature. As modularity is not the stable end state of industry evolution, firms must retain their system-level knowledge. It is again Chesbrough (2003b) who pushes this analysis one step further. He asks how firms can develop their system-level knowledge in a modular world, when they must compete against firms that pursue highly focused business strategies, and thus do not incur the cost of developing and maintaining the systems-level knowledge? In his words, "(o)nce the advance of modularity has obliterated the rents from architectural knowledge, there is a real question of how a firm may sustain its systems integration capabilities in the absence of being able to profit from that knowledge." (ibid: : 191).

## 6. Responses

Let us finally briefly examine the difficulties that management faces when it attempts to respond to the limits to modularity described in this paper. A useful analytical device is the concept of "iterated co-design", introduced by Sabel and Zeitlin (2004). Defined as adjustments in organization, work routines and location that allow for formalized but flexible inter-firm collaboration within global networks, this concept is particularly helpful to understand implications for the organization and strategies of engineering and design networks.

A basic characteristic of "iterated co-design" arrangements is that "suppliers contribute to the redefinition of interface specifications for new products based on their experience in manufacturing existing models." (Sabel and Zeitlin, 2004: 12). As a decentralized process of concurrent engineering, "iterated co-design" has been pioneered in the car industry. But, as research on global design networks in the electronics industry shows, SoC design methodologies are following similar principles (Ernst, 2004b). Iterated co-design has two objectives: to cope with uncertainty that incumbent market leaders face through disruptive technologies, as defined in Christensen (1997); and to shape and adjust the multiple interfaces among disintegrated and geographically dispersed design teams that are contributing to a modular design project.

"Disruptive technologies" are a serious challenge for incumbent market leaders, precisely because they begin to realize their potential in lower-end and initially peripheral markets, and because they are typically cheaper, simpler, smaller, and, frequently, more convenient to use. (Christensen, 1997: XV) Christensen provides three reasons why incumbents fail to notice "lower-end" markets that may erode their market leadership: i) Because these technologies are simpler and cheaper, they promise lower margins, not greater profits: "It is very difficult for a company whose cost structure is tailored to compete in high-end markets to be profitable in low-end markets as well" (Christensen, 1997, p.XX). ii) Disruptive technologies are first commercialized in emerging and insignificant markets that large companies have great difficulties to address. And iii) The incumbents' most profitable customers generally do not want, and initially cannot use products based on disruptive technologies.

"Disruptive technologies" are ubiquitous, because, in principle, all established technologies are disruptible. This explains why "modularization, by entrenching deep assumptions about the design of products or product classes, is a self-limiting strategy." (Sabel and Zeitlin, 2004:8). With high uncertainty about changes in technology and markets, stabilization of technical interface standards is a two-edged sword. "Some standardization is obviously necessary to allow specialists to focus on the complex subsystems in which they have distinctive capabilities. But too much standardization can just as obviously become a barrier to systematic innovation and lock component manufacturers into a potentially obsolete product architecture." (Sabel and Zeitlin, 2004: 9).This explains why firms, especially in fast-moving industries like electronics, are reluctant to commit to a single product architecture and to a single set of technical interface standards, hence constraining the development of market modularity.

Of equal importance however is the second objective of "iterated co-design", i.e. to coordinate the multiple interfaces that reflect the growing complexity of SoC design. An interface is created when information must flow, and when knowledge must be exchanged, "…between groups that are isolated from each other, whether by goals, methodologies, geography or culture" (Wilson, 2003: 49). This is exactly what is happening with SoC design. The diversity of functions that must be integrated into the chip means that "various blocks within the finished design will have come from different groups, some within and some outside the design team. Some of these groups …may not share a vocabulary, or even a language and culture with the primary chip design group." (Wilson, 2003:48). A typical SoC design team needs to manage at least six main types of design interfaces (Ernst, 2004b): with system designers, with providers of design building-blocks (the so-called silicon intellectual property, SIP), with software developers, with verification teams, with EDA tool vendors, as well as with foundry services (fabrication).

In addition, these different design network communities are spread across the ten strategic groups of firms and research institutes that participate in GDNs, that I have described in note 3 of this paper. Managing these multiple design interfaces poses extremely demanding coordination requirements. For instance, each of the different design network communities insists on using their own language and tools. Typically, in SoC design teams, there are "islands of automation" of different design tasks, each based on a different language. "Nobody will easily give up the language or the approach used for their own particular task for the sake of the overall flow" (Pierre Bricaud, design manager at Mentor Graphics, quoted in "SoC designers describe their 'best practices'", 2002).

This poses a serious challenge to modular design organization. As design teams become larger and separated by distance or design disciplines, they need to be able to communicate with each other. While they share a common objective, they use highly dissimilar vocabularies. Defining interfaces requires shared definitions of the data that need to be exchanged, of the formats and protocols that govern data transfer and interpretation, and of the economic performance requirements of the designs. Developing a precise common vocabulary for these three interface attributes is extremely difficult. Equally important, data must be translated into a form usable by different design groups.

"Iterated co-design" is an attempt to cope with these difficulties, by rendering tacit knowledge at least partly explicit. But this requires a profound capacity for systems integration in order to coordinate the multiple interfaces between. Equally important is a capacity for close monitoring of interactions, in order "to detect performance failures and deception before they lead to disastrous consequences" (Sabel and Zeitlin, 2004: 13). To understand the challenges for "iterated co-design", let us look at two illustrative examples of design interface management within GDNs: interfaces of the core SoC design team with system designers, and interfaces with foundry services.

By definition, SoC design requires close interaction with system designers, marketing people and end customers ( the "set makers"). With product life cycles often as short as six months or less, system design requirements keep changing rapidly. The protocol necessary to transmit these changes real-time to all the different design network participants is "one of the great unsolved problems of design management". (Wilson, 2003: 56). Communication is also fraught with problems between hardware and software designers: "Verilog representation of a block by hardware designers, for example, may need to be translated into …. a Java transaction-level model of the block for use by software designers." (Wilson, 2003: 52). The amount of work required for this translation is enormous, so much depends on the availability of new software tools that would make a progressive automation possible.

Interactions with foundry services are arguably the most explicitly recognized interfaces in the entire SoC flow, with well documented and automatically checkable "design rules" (Macher, Mowery and Simcoe, 2002). Yet, with growing complexity of SoC design, the management of the foundry interface also poses new challenges. A combination of new processes and drastic changes in design methodology implies that design rules need to be tweaked and stretched, and that process limitations are there "to be explored not worshipped" (Wilson, 2003: 63), requiring a much closer interaction between designers and process engineers. From the perspective of foundries, for instance, sub-micron process technology makes yield enhancement much more difficult for 20 to 50 million transistor SoCs (author's interview with Taiwanese foundry, September 22, 2003).

As processes grow more demanding, mask makers and process engineers will try to pass this growing complexity to the chip design team through an enormous increase in complexity in either cell selection or design rules. For instance, a leading-edge SoC design is likely to require 22 separate cells. And the foundry's process engineers now include into the design rules for SoC designers the request to "design-for-yield-enhancement."

This new interface requirement with fabrication means that design teams must adjust the design to improve the odds that the process will yield well and that the dice will continue working even under demanding system performance requirements. In other words, designers must take into account the effects of fabrication process variations, which makes design even more complex. There is now a much greater need for dense interaction between physical designers and process integration teams, even for relatively stable designs. Designers increasingly must take into account the intricacies of process development. An "extraordinary degree of coordination" is required between SoC designers, mask makers, foundries, and third party SIP suppliers ("90-nm design flow is seen as a community effort", 2003).

In short, chip design has become itself a highly complex technology system, where multiple communication and knowledge exchange interfaces must be managed simultaneously. Obviously, the idea of translating technical modularity into organizational modularity through vertical specialization has many attractions. Yet, its implementation requires a mind-boggling degree of cooperation among the diverse participants of design networks. We have seen that this is true for all the different design interfaces. Ironically, the more chip design becomes modularized through SoC methodologies, and the more it moves into sub-micron territory, the more current transformations appear to move in the opposite direction to more integrated forms of industrial organization. Yet this does not imply a return to the *status quo ante* of vertical integration. Instead, new hybrid forms of industrial organization are likely to emerge, as firms need to cope with increasingly complex technical and competitive challenges (as explored in Ernst and Martin, 2004).

## Conclusions

To conclude, it is time to revisit earlier debates on modularity in order to improve the theoretical and empirical validity of that concept. Progress in the division of labor in design (technical modularity) has obviously created ample opportunities for the organization of firms beyond vertical integration. It may also transform industry structure and markets, by fostering vertical specialization in both manufacturing and knowledge production. But there is nothing automatic about this process. That we do not find a "natural" convergence of technical, organizational and market modularity has its root cause in three fundamental limits to modularity that the paper has explored in detail, drawing on evidence from chip design: demanding coordination requirements, especially for technology development; constraints to interface standardization; and a slow-down of innovation that may result from too strict an adherence to market and organizational modularity.

Drawing on a revisionist literature that has emerged over the last few years, the paper has highlighted three specific weaknesses of current theories of modularity: they confound technical with market modularity; they exaggerate the empirical significance of market modularity; and they fail to explore the dynamics of modularization. The paper argues that these weaknesses must be addressed, if we want to capture the complex forces and constraints that link technical modularity with changes in organization and market structure.

All of this strongly suggests that we need to counter whatever tendency exists in the "modularity" literature to generalize empirical observations that are context-specific and to confound them with prescription as well as prediction. It is certainly not possible to sustain the claim that modularity is the stable end state of industry evolution, and that this is true across industries and technologies. While modular design has acted as a powerful catalyst for changes in business organization and industry structure, limits to modularity are aplenty, and constrain the convergence of technical, organizational and market modularity.

To conclude, these findings add an important qualification to a widely accepted proposition that, as relevant knowledge can now be codified in frameworks and categories, codification will "naturally" enhance the division of innovative labor by lowering the transaction cost of technological knowledge (e.g., Arora and Gambardella, 1994). As Tokumaru (2004: 3) observes, this concept of "knowledge" comes very close to Arrow's (1962) concept of context-independent information that can be easily transmitted. By reducing learning to information processing, that proposition neglects the increasing complexity of technological knowledge that accompanies the process of codification, as this paper demonstrates for chip design.

The paper also adds a second important qualification to the study of internationalization of innovation. I emphasize that vertical specialization does not imply that the "Visible Hand" of large global corporations will become invisible, giving rise to a resurgence of market forces. In line with Pavitt (2003a and 2003 b), Chesbrough (2003a and 2003b), Brusoni (2003), Tokumaru (2004) , and Brusoni, Prencipe and Pavitt (2001), this paper introduces a counter-proposition: if codification does not reduce complexity (which it fails to do in chip design), then the division of innovative labor will remain constrained, and requires more (not less) coordination through system integrators or network flagships.

## References

Acha, V., 2002, Framing the Past and Future: The Development and Deployment of Technological Capabilities in the Upstream Petroleum Industry, unpublished PhD thesis, Science Policy Research Unit (SPRU), University of Sussex at Brighton

Acha, V. and S. Brusoni, 2002, "Knowledge on Wheels: New Lessons from the Tyre Industry", paper presented at the June 2002 DRUID Conference, Copenhagen, Denmark

Arora, A., and A. Gambardella, 1994, "The Changing Technology of technological Change: General and Abstract Knowledge and the Division of Innovative Labor", Research Policy, 23: 523-532

Arora, A., A. Gambardella and A. Rullani, 1998, "Division of Labour and the Locus of Inventive Acitivity", Journal of Management and Governance, 1: 123-140

Arrow, K.J., 1962, "The Economic Implications of Learning by Doing", Review of Economic Studies, June

Baldwin, C.W. and K.B.Clark, 2000, Design Rules: The Power of Modularity, MIT Press, Cambridge, Mass

Bengtsson, L. and C. Berggren, 2002, "Horizontally Integrated or Vertically Divided? A Comparison of Outsourcing Strategies at Ericsson and Nokia", Working Paper, Linkoeping University, Sweden, October

Benson-Armer, R.J., D.L. Dean, and J.C. Kelleher, 2004, "Getting Contract Manufacturers Back on Track", The McKinsey Quarterly, July

Brusoni, S. and A. Prencipe, 2001, "Technologies, Product, Organisations: Opening the Blackbox of Modularity", Industrial and Corporate Change, Vol. 10(1):179-205

Brusoni, S., 2003, "Authority in the Age of Modularity", SPRU Electronic Working Paper Series, No. 101, The Freeman Centre, University of Sussex, June

Brusoni, S., A. Prencipe and K. Pavitt, 2001, "Knowledge Specialisation, Organizsational Coupling, and the Boundaries of the Firm: Why Do Firms Know More Than They Make?", Administrative Science Quarterly, 46: 597-621

Chandler, A. D., 1977, The Visible Hand. The Managerial Revolution in American Business, The Belknap Press of Harvard University Press, Cambridge, Massachusetts

Chang, H. , L. Cooke, M. Hunt, G. Martin, A. McNelly and L. Todd, 1999, Surviving the SoC Revolution: A Guide to Platform-Based Design, Kluwer Academic Publishers, Boston, etc

Chesbrough, H.W., 2003a, Open Innovation. The New Imperative for Creating and Profiting from Technology, Harvard Business School Press

Chesbrough,. H.W., 2003b, “Towards a Dynamics of Modularity. A Cyclical Model of Technical Advance”, in: Prencipe, A., A. Davies and M. Hobday, eds, The Business of Systems Integration, Oxford: Oxford University Press

Chesbrough, H. and K. Kusunoki, 2001, “The Modularity Trap: Innovation, Technology Phases Shifts and the Resulting Limits of Virtual Organizations”, in: I. Nonaka and D. Teece (eds.), Managing Industrial Knowledge, London: Sage Press, 202-30

Davies, A., 1997, “The Life Cycle of a Complex Product System”, International Journal of Innovation Management, Vol.1, #3: 229-256

Davies, A., 2003, “Integrated Solutions. The Changing Business of Systems Integration”, in: Prencipe, A., A. Davies and M. Hobday, eds, The Business of Systems Integration, Oxford: Oxford University Press

Ernst, D., 1983, , The Global Race in Microelectronics, with a foreword by Prof. David Noble, Massachusetts Institute of Technology (MIT), Campus Publishers, Frankfurt and New York, 290 pages

Ernst, D., 2002a , “Global Production Networks and the Changing Geography of Innovation Systems. Implications for Developing Countries”, Journal of the Economics of Innovation and New Technologies , Vol. XI, #6, pp. 497-523

Ernst, D., 2002b, “The Economics of Electronics Industry: Competitive Dynamics and Industrial Organization”, in: The International Encyclopedia of Business and Management (IEBM), Handbook of Economics, editor: William Lazonick, International Thomson Business Press, London

Ernst, D., 2003, “Digital Information Systems and Global Flagship Networks: How Mobile is Knowledge in the Global Network Economy?”, in: J.F. Christensen (ed.), The Industrial Dynamics of the New Digital Economy, Edward Elgar, Cheltenham

Ernst, D., 2004 a, “Global Production Networks in East Asia’s Electronics Industry and Upgrading Perspectives in Malaysia”, in: S. Yusuf, M.A. Altaf and K. Nabeshima (eds.), Global Production Networking and Technological Change in East Asia, a co-publication of the World Bank and Oxford University Press, Washington, D.C.

Ernst, D., 2004b, “Internationalisation of Innovation: Why is Chip Design Moving to Asia?”, in International Journal of Innovation Management, special issue in honour of Keith Pavitt (Peter Augsdoerfer, Jonathan Sapsed and James Utterback, guest editors); also available as East-West Center Economics Working Paper # 64

Ernst, D., 2004c, “The New Mobility of Knowledge: Digital Information Systems and Global Flagship Networks”, in: S. Sassen (ed.), Digital Formations in a Connected World, published for the U.S. Social Science Research Council, Princeton University Press, also available as East-West Center Economics Working Papers, No. 56, June

Ernst, D. and O'Connor, D., 1992, Competing in the Electronics Industry. The Experience of Newly Industrialising Economies, Development Centre Studies, OECD, Paris, 303 pages

Ernst, D. and G. Martin, 2004, “Co-Evolution of Technical Modularity and Organizational Integration - the Case of “System-on-Chip” Design”, manuscript, East-West-Center, Honolulu

Flamm, K. (1999), ‘Digital Convergence? The Set-Top Box and the Network Computer’, in J.A. Eisenach and T.M. Lenard (eds), Competition, Innovation and the Microsoft Monopoly: Antitrust in the Digital Market Place, Boston: Kluwer Academic Publishers

Foray, D. and Steinmueller, W.E., 2001, ‘Replication of routine, the domestication of tacit knowledge and the economics of inscription technology: a Brave New World?’, paper presented at the conference in honor of Richard R. Nelson and Sidney Winter, Danish Research Unit in Industrial Dynamics (DRUID), Aalborg, Denmark, June 12-15

Gawer, A. and M.A. Cusumano, 2002, Platform Leadership. How Intel, Microsoft and Cisco Drive Industry Innovation, Harvard Business school Press, Boston, Mass.

Gereffi, G, 1994, “The Organization of Buyer-Driven Global Commodity Chains: How U.S. Retailers Shape Global Production Networks”, in: G. Gereffi and M. Korzeniewicz, eds., Commodity Chains and Global Capitalism, Greenwood, Westport, Conn, pages 95-122

Granstrand, O., P. Patel and K. Pavitt, 1997, “Multi-Technology Corporations: Why They Have “Distributed” Rather Than “Distinctive Core” Competencies”, California Management Review, Vol. 39 #4, pp. 8-25

Henderson,R.M. and K.B.Clark, 1990, “Architectural Innovation: The Reconfiguration of Existing Systems and the Failure of Established Firms”, Administrative Science Quarterly, March, 9-30

Hobday, M., 1995, Innovation in East Asia: The Challenge to Japan, Edward Elgar, Aldershot

James, G., 2004, “Panacea or Snake Oil?”, Electronic Business, 1 June

Jørgensen, H.D. and J. Krogstie (2000), ‘Active Models for Dynamic Networked Organisations’, Working Paper, Institute of Computer & Information Sciences, Norwegian University of Science and Technology, available from Harvard.D.Jorgensen@ informatics.sintef.no

Kogut, B. and U. Zander (1993), ‘Knowledge of the Firm and the Evolutionary Theory of the Multinational Corporation’, Journal of International Business Studies, 24 (4), 625.

Lakenan, B., D. Boyd and E. Frey (2001), ‘Why Cisco fell: Outsourcing and its Perils’, Strategy + Business, Booz-Allen &Hamilton, Issue 24, 3rd Quarter, 18 -32.

Langlois, R.N., 2003, “The Vanishing Hand: The Changing Dynamics of Industrial Capitalism”, Industrial and Corporate Change, Vol. 12, #2, pp.351 -385

Linden, G. and D. Somaya, 2003, “System-on-a-chip integration in the semiconductor industry: industry structure and firm strategies”, Industrial and Corporate Change, Volume 12, # 3: 545-576

Luethje, B, 2002, “Electronics Contract Manufacturing: Global Production and the International Division of Labor in the Age of the Internet”, Industry and Innovation, special issue “Global Production Networks”, volume 9, no. 3, editors: Dieter Ernst and Linsu Kim

Macher, J.T., D.C. Mowery, and T.S. Simcoe (2002), ‘eBusiness and the Semiconductor Industry value Chain: Implications for Vertical Specialization and

Integrated Semiconductor Manufacturers', Industry and Innovation, special issue 'Global Production Networks', 9 (3), guest editors: D. Ernst and L. Kim, 155-182

Martin, G. and H. Chang, eds., 2003, Winning the SoC Revolution: Experiences in Real Design, Kluwer Academic Publishers, Boston etc

Pavitt, K., 1999, Technology, Management and Systems of Innovation, Edward Elgar, Cheltenham

Pavitt, K., 2003a, "Are Systems Designers & Integrators 'Post-Industrial' Firms?'", in A. Prencipe, A. Davies and M. Hobday (eds), Systems Integration and Firm Capabilities, Oxford: Oxford University Press

Pavitt, K., 2003b, "What are Advances in Knowledge Doing to the Large Industrial Firm in the 'New Economy'?" in Jens F. Christensen and Peter Maskell, eds., The Industrial Dynamics of the New Digital Economy, Edward Elgar, Cheltenham etc

Pisano, 1996

Prencipe, A., 1997, "Technological Capabilities and Product Evolutionary Dynamics: a Case Study from the Aero Engine Industry", Research Policy, 25: 1261-1276

Prencipe, A., A. Davies and M. Hobday, (eds), 2003, The Business of Systems Integration, Oxford: Oxford University Press

Sabel, C. F. and J. Zeitlin, 2004, "Neither Modularity nor Relational Contracting: Inter-Firm Collaboration in the New Economy", *Enterprise and Society*, 5, 3 (September)

Sichel, D.E. (1997), The Computer Revolution: An Economic Perspective, Washington: Brookings Institution Press.

Simon, H., 1962, "The Architecture of Complexity", Proceedings of the American Philosophical Society, 106,6:467-82

"SoC designers describe their 'best practices'", editorial, *EEdesign*, April 5, 2002

Sturgeon, T., 2002, "Modular Production Networks: A New American Model of Industrial Organization", Industrial and Corporate Change, Vol. 11/#3

Takeishi, A. and T. Fujimoto, 2003, "Modularization in the Car Industry: Interlinked Multiple Hierarchies of Product, Production and Supplier Systems", in: Prencipe, A., A. Davies and M. Hobday, eds, The Business of Systems Integration, Oxford: Oxford University Press

Tokumaru, Norio, 2004, "Codification of Technological Knowledge, Technological Complexity, and Division of Innovative Labour", in J.H. Finch and M. Orrillard, eds., Complexity and the Economy: Implications for Economic Policy, Edward Elgar

von Hippel, E, 1990, "Task Partitioning: An Innovation Process Variable", Research Policy, 19: 407-18

Wilson, R, 2003, "Non-Technical Issues in SOC Design", Martin, G. and H. Chang, eds., Winning the SoC Revolution: Experiences in Real Design, Kluwer Academic Publishers, Boston etc