Analyzing the Effects of Product Architecture on Technical Communication in Product Development Organizations

by

Manuel E. Sosa R.

Ingeniero Mecánico
Universidad Simón Bolívar, 1992

Master of Science in Mechanical Engineering
Massachusetts Institute of Technology, 1996

Submitted to the Department of Mechanical Engineering
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY IN MECHANICAL ENGINEERING
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2000

Signature of Author: ______________________
Manuel E. Sosa R., Department of Mechanical Engineering
May 22, 2000

Certified by: ______________________
Steven D. Eppinger, Associate Professor of Management
Thesis Supervisor, Doctoral Committee

Certified by: ______________________
David R.Wallace, Associate Professor of Mechanical Engineering
Chairman, Doctoral Committee

Certified by: ______________________
Thomas J. Allen, Professor of Management
Doctoral Committee

Accepted by: ______________________
Ain A. Sonin, Professor of Mechanical Engineering
Chairman, Committee on Graduate Students

Analyzing the Effects of Product Architecture on Technical Communication in Product Development Organizations

by

Manuel E. Sosa R.

Submitted to the Department of Mechanical Engineering
on May 22, 2000 in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Mechanical Engineering

Effective communication in product development organizations has been identified as a key factor of product development performance. Furthermore, understanding how the development organization manages the knowledge associated with the product architecture is broadly recognized as a critical challenge for established firms facing architectural innovation. This thesis presents a research method and statistical analyses intended to enhance understanding of the coupling between the product architecture and the development organization.

The research method is summarized by three steps: 1) capture the product architecture by documenting design interfaces, 2) capture the development organization by documenting team interactions, and 3) couple the product architecture with the development organization by comparing design interfaces with team interactions. Our approach is illustrated by analyzing the development of a large commercial aircraft engine.

Several hypotheses are formulated to explain the mismatch between design interfaces and team interactions, that is, the cases when: 1) known design interfaces are not matched by team interactions, and 2) reported team interactions are not predicted by design interfaces. Effects due to organizational and system boundaries, design interface strength, design interface type, design interface redesign, indirect team interactions, and secondary design interfaces are studied. In addition, through the analysis of the distribution of cross-boundary design interfaces, modular and integrative systems are formally identified, and differences between designing modular versus integrative systems are studied.

Two types of statistical analyses were performed. First, categorical data analysis techniques were used to test the mentioned hypothesized effects. Second, a log-linear model, built upon social network analysis methods, was developed to study the association between design interfaces and team interactions controlling for effects of reciprocity, differential attraction, and differential expansiveness of both components and teams. Findings in this research are complemented with the results of another empirical study focused on the effects of distance and communication media use in geographically distributed development organizations.

By considering the results presented in this thesis, development organizations can improve the integration process for complex designs. The approach developed is particularly applicable to projects where the architecture of the product is well understood and the development team is organized around the product architecture.

Thesis committee: Prof. Steven D. Eppinger (thesis supervisor)
Prof. David R. Wallace (chairman)
Prof. Thomas J. Allen

Biographical Note

Manuel Emilio Sosa Rosales

(*msosa@alum.mit.edu*)

22 January 1971 Born in Mérida, Venezuela.

Education

1987-1992 Universidad Simón Bolívar. Caracas. Venezuela.
Bachelor of Science degree in Mechanical Engineering. *CUM LAUDE.*

1994-1996 Massachusetts Institute of Technology.
Master of Science degree in Mechanical Engineering
Research area: Design Optimization and Automation.

1996-2000 Massachusetts Institute of Technology
Ph.D. degree in Mechanical Engineering
Major: Product Development. Minor: Business Dynamics

Professional Experience

1992-1994 Petrochemical of Venezuela (PEQUIVEN). Morón, Venezuela.
Systems Engineer.

Summer 1995 Mitsubishi Electric Research Laboratories. Cambridge, MA.
Summer Researcher.

Summer 1996 Parametric Technology Corporation. Waltham, MA.
Technical Marketing Associate.

Research Interests Information technology in product development;
Business dynamics; Network analysis methods.

Written Work (http://cadlab.mit.edu/users/msosa/research_phd.html):

Sosa, M. E. "Optimizing the Packing of Three-Dimensional Shapes Using Genetic Algorithms", *unpublished Master's thesis in Mechanical Engineering*, MIT, 1996.

Sosa, M. E., Eppinger, S. D., and Rowles, C. M. "Designing Modular and Integrative Systems", *Proceedings of ASME DTM Conference,* DETC2000/DTM14571, September 2000. (Nominated for Best Paper Award)

Sosa, M. E., Eppinger, S. D., and Rowles, C. M. "Understanding the Effects of Product Architecture on Technical Communication in Product Development Organizations".

Sosa, M. E., Eppinger, S. D., and Pich, M. "Factors that Influence Technical Communication in Distributed Product Development: An Empirical Study in the Telecommunications Industry".

Acknowledgments

Wow, and now I am writing the last portion of my thesis. This is perhaps the most important part of the thesis because it makes me realize about the ups and downs of the last four years, and the people who has directly or indirectly contributed on my achieving this goal.

I still remember the day that I fully decided to pursue a doctoral degree. I was making the conscious decision of formally learning how to identify relevant problems and seek for innovative ways of solving them. To achieve such a demanding goal I put great emphasis on choosing the person who would be my mentor. Well, I cannot express in words how fortunate I have been of choosing Prof. Steven D. Eppinger as my advisor. Steve has surpassed all positive expectations I had for what I had considered the ideal advisor was. He has made the doctoral program the most fruitful learning experience of my life. I have learned a new lesson, in many aspects of life, from every single interaction we have had all these years. Steve has timely provided me with the right doses of freedom, demand, encouragement, and wisdom. It has been great to have a mentor who believes in one as a student, as a researcher, and as a person.

The members of my thesis committee have been tremendously helpful throughout this thesis. Having Dave Wallace available in an everyday basis has provided me with a great deal of informal communication essential to succeed in a research project of this magnitude. Dave never forced the agenda on me, but instead contributed important guidance to pursue my own path. Having Tom Allen, one of the pioneers in this research area, as one of my thesis committee members has been a thrilling idea even before he accepted to be part of this team. I am very grateful to him for accepting being part of this committee, and even more grateful for the enormous contributions of his advises, questions, comments, and suggestions throughout this project.

This work would have been almost impossible to complete without the great help of Craig Rowles, a former SDM student, on the data collection. Our discussions to learn as much as we could from the data collected were very insightful. I am also pleased to thank Christophe Van den Bulte, from The Wharton School, for his patience and willingness to share with me all his knowledge about social network analysis. It would have been impossible for me to learn so much so fast on that topic without his help.

The main reason we have the sixth chapter on this thesis is due to professors Michael Pich and Christophe Loch, from INSEAD. I am in great debt with both of them for hosting me at INSEAD during the Summer of 1998. That was one of the most productive summers of the last years. The discussions with Christophe were very insightful. I am very grateful to Mike for allowing me to explore his database, and encourage a joint research effort that has led to great results. The summer of 1998 would have not been such an amazing experience without the friendship of the doctoral students at INSEAD. Thanks to Stelios, Sezer, Goncalo, Nicola, and Stephanie for made me feel part of the group while I was there. I am particularly grateful to Stelios for his unconditional friendship and his willingness to read and discuss my crude paper drafts. His comments have been of great benefit in achieving a better understanding of my thoughts. Also, thanks to Paolo for being such a good host in Italy.

I cannot forget the people who gave me invaluable support during the first years of the program. My conversations with Prof. Ernesto Blanco were incredibly helpful to gain the maturity needed to face the challenge that it was going about to start. Thanks to other doctoral students Tim, Jim, Gu, Siduh, Jean Pierre, Osamah, Asif, and Jose, with whom I shared the struggling feelings of the first years.

I also want to thank other cadlabbers from whom I learned many lessons. I miss my conversations with Rahul, Krish, Kazu, Jay, Alok, Francis, Jun-Beom, Chun-ho, Dave, Paul, Naren, Shaun, Ashley, Nick, and Nicola. Thankfully the quality of cadlabbers has not decreased with time and I have been able to have good discussions with current cadlabbers Shaun, Ines, Priscilla, Bill, Stephen, Jeff, Juan Carlos, Jaehyun, Julie, Maria, and the rest of the gang.

Our DSM research team has been great. I am particularly grateful to Maria for all the quality time we spent together talking about research and life. I also thank Tyson, Nader, Inaki, Vesas, Jen, Soo-Haeng, Ali, and Dan for productive research conversations. I also thank Cara for being such a good assistant.

I am very pleased and proud of being part of the Center for Innovation in Product Development (CIPD) which has provided the funding to support this research effort. The existence of the CIPD is clear evidence of how important interdisciplinary research is. I thank the directors, professors, students and staff of the CIPD for their constant support. I am also grateful to Leslie and Joan for their support and their smiles every time I dropped by the mechanical engineering graduate studies office.

Many people have been indirectly involved in this project just by keeping me busy with other activities. I want to thank the people I have lived with all these years, Alvarito, Ingrid, Nicolas, and Chris, they helped a great deal just for being there every day. Thanks to my unconditional friends Mark, Laura and Dom, for the extra brainpower they have given me since we met. Special thanks to Tony for being such an outstanding friend since I arrived to MIT, and for his perseverance on keeping me busy having fun.

There are many venezuelans I want to thank for their friendship and support, but then this section would be even longer. Special thanks to Enrique and Carlos for such a great system dynamics project, and to Leonela for her positive energy. Thanks to my "super-panas" in Venezuela Vitolo, Fofo, and Serryo whom have been always in touch all these years.

I have been incredibly privileged of having a wonderful host family. The Briggss have contributed a great deal to my experience of living in US. Dave and Elaine have been a source of constant support, happiness, advises, and good experiences. I just love them with all my heart.

I want to thank "mi chamita" for sharing with me a wonderful time in the Boston area, for her unique way of enjoying life, and for teaching me how important listening to others is.

I can state with all certainty that this thesis would have never been completed without the support of all my family. Special thanks to my aunts Milagros and Morella. A million thanks to "abuelita" for her infinite prayer. I thank my sister, Ana Beatriz, for spending so much time listening to me, even when I am senseless. She has been an infinite source of encouragement to me. I have no words to describe how thankful I am to my mom. She is without any doubt the most amazing person in this planet. As a friend, as a professional, as mom (and dad), she has been the greatest. She deserves all the credit for anything good I do.

Lastly, but certainly not least, I have to thank God for providing me with the stamina to start and conclude this project, and for putting all the people I mentioned above on my path through life.

Abuelita, linda, and mami:
This is for you all!

Table of Contents

List of Figures

List of Tables

1. *Introduction*

"Hay hombres que luchan un día
y son buenos.
Hay otros que luchan un año
y son mejores.
Hay quienes luchan muchos *años*
y son muy buenos.
Pero hay los que luchan toda la vida:
esos son los imprescindibles."
-Bertolt Brescht
(en Sueño con Serpientes (1975), de Silvio Rodríguez)

"There are men that fight for a day
and they are good.
There are others that fight for a year
ond those are better.
There are those that fight for many years
and those are very good.
But there are those that fight all their lives:
those are the indispensable."
-Bertolt Brescht
(in I Dream about Snakes (1975), by Silvio Rodríguez)

The increasing need to compete in established markets as well as to address new markets in order to sustain corporate growth is adding more pressure onto product development organizations to improve their development performance. This thesis introduces a method to understand to what extent the architecture of a product determines the technical interactions between design teams.

This work is motivated by the crucial importance of product development in today's businesses and the need to improve our understanding of the communication process in development organizations. Much has been written about process improvement in the product development arena and in particular about the role of effective communication in product development teams. Allen (1977) pioneered the stream of research dedicated to investigate how effective internal and external communications stimulate the performance of development organizations. Clark and Fujimoto (1991) relate successful development in the auto industry to intensive communication between upstream and downstream activities. Wheelwright and Clark (1992) emphasize the need to improve communication when and where it certainly improves project performance. Ulrich and Eppinger (1995) also emphasize the need to facilitate the exchange of essential information in order to speed up the development process.

Under the information processing perspective introduced by Alexander (1964), a product development process transforms a set of inputs (e.g. customer needs, product strategy, manufacturing constraints) into a set of outputs (e.g. product design). This typically requires that members of a product development team communicate with others, either within or outside the development team, in order to accomplish their development activities. Thus, communication becomes an important correlate of R&D performance (Allen (1964); Keller (1986); De Meyer (1991); Håkanson and Nobel (1993)). As De Meyer noted, "one of the most important productivity problems in R&D is stimulating communication among researchers" (1991: p. 49).

The objective of our study is to analyze the coupling of the architecture of the product to be developed and the structure of its product development organization. By predicting technical communication in this way, we aim to provide a method that improves planning of development projects where the architecture of the product is known in advance.

From a strategy viewpoint, Henderson and Clark (1990) identified how critical it is for established firms to recognize novel product architectures. Furthermore, improved development of architectural knowledge provides a competitive advantage for firms facing architectural innovation. The approach illustrated in this thesis provides important insights for managers addressing product and organizational changes.

1.1. Complex Product Development Leads to Decomposition and Integration

This thesis addresses the problem of understanding technical communication in complex product development. We focus on the development of complex but relatively mature products, such as an automobile, a computer, or an aircraft engine. The general approach when developing complex products is to decompose the product into systems, and if the systems are still too complex, decompose these into smaller sets of components (Alexander (1964), Simon (1981), Smith and Browne (1993), McCord and Eppinger (1993), Pimmler and Eppinger (1994), Eppinger (1997)). Consequently, product architecture is defined as the scheme by which decomposed elements of a product are arranged into sets of components in order to meet its functional requirements (Ulrich and Eppinger, 1995).

Ulrich (1995) defines several types of architectures according to how the product's functions are mapped onto its physical components. A key feature of product architecture is the degree to which it is modular or integral. Modular architectures exhibit direct mapping between functions and physical elements, and have well-defined interfaces between physical components. On the other hand, integral architectures spread functions across physical components, resulting in more complex interfaces between them (Ulrich and Eppinger, 1995). In very complex products, we apply these definitions at the level of the many systems (and subsystems) which comprise the product. We will refer to *modular systems* as those exhibiting modular architecture characteristics while *integrative systems* are those revealing integral architecture features.

From an organizational standpoint, teams are commonly organized around the architecture of the product. In most technical products we can observe a clear mapping between the product architecture and the development organization which designs it (McCord and Eppinger (1993), Pimmler and Eppinger (1994)). Complex development projects usually involve the efforts of hundreds or even thousands of team members. A single team does not design the entire product at once (too complex). Rather, many teams develop the components, or systems, and work to integrate all of these components to create the final product (von Hippel (1990)).

Design teams face two important levels of integration during the development of complex products. Function-level integration takes place within each cross-functional design team when they have to coordinate efforts in order to design their respective components. System-level integration takes place across design teams in order to integrate the components (designed by each team) to assure the product works as an integrated whole. Furthermore, we distinguish two types of system-level integration efforts. Within-group system-level integration effort, which takes place between teams that design components of the same system. Across-group system-level integration effort, which takes place between teams that design components that belong to different systems.

Figure 1.1 illustrates the various levels of integration faced by the development organization. This thesis focuses on system-level integration efforts. We aim to better understand how the coupling of product architecture and organizational structure drives system-level coordination efforts across design teams.

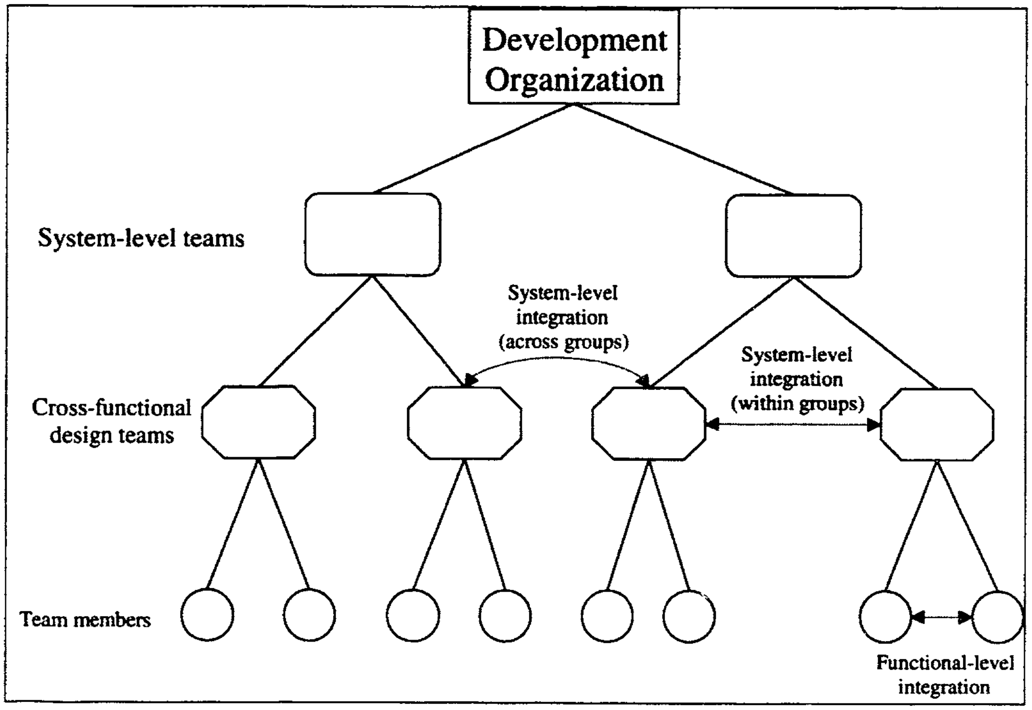

Figure 1-1. Levels of Integration Effort in a Development Organization

1.2. Technical Communication in Development Organizations

Previous research (Allen (1997), Morelli *et al* (1995)) has identified three types of technical communication in development organizations.

- **Coordination-type**. Team members communicate to coordinate their tasks. Technical information transfer about task related issues. This type of communication is critical in project-based organizations as the one we are studying on this paper. In this type of organization better project performance is achieved by accomplishing effective communication across disciplines (Allen and Hauptman (1990)).
- **Knowledge-type**. Team members communicate with their peers to keep up to date with the latest developments in their disciplines. Consultation, instruction and skill development. This type of communication is particularly important to maintain technology currency within specific disciplines, which is the underline goal of functional organizations (Allen and Hauptman (1990)).
- **Inspiration-type**. Team members communicate for creativity, inspiration, and managerial affirmation. Motivation of team members.

For the purpose of this thesis, we are concerned with coordination-type communications, which are directly related to the system-level integration effort across design teams we want to understand. Several researchers have focused their efforts on predicting and understanding communication patterns in development organizations. Allen (1977, 1997) has proposed models based on distance separation between team members in R&D organizations to estimate the probability they engage in technical communication. Griffin and Hauser (1992) showed that using Quality Function Deployment (QFD) practices enhances technical communication within the boundaries of the development teams, but reduces the communication levels across teams' boundaries. Morelli *et al.* (1995) showed that coordination-type communications could be predicted by analyzing the task structure of development projects. Van den Bulte and Moenaert (1998) studied the effects of R&D collocation on cross-functional communication. Finally, we present in chapter seven an empirical study in the telecommunications industry about how distributed development organizations use various communication media (Sosa *et al.* (2000)).

1.3. Research Questions

Complex products are decomposed into systems, and these systems are further decomposed into components. The arrangement of these physical sets of components defines the architecture of the product. Similarly, development organizations are usually split into design teams that develop each of the components that comprise the product. Figure 1.2 illustrates the main research question we want to investigate: How does the architecture of a product drive the technical communications between the teams which design it?

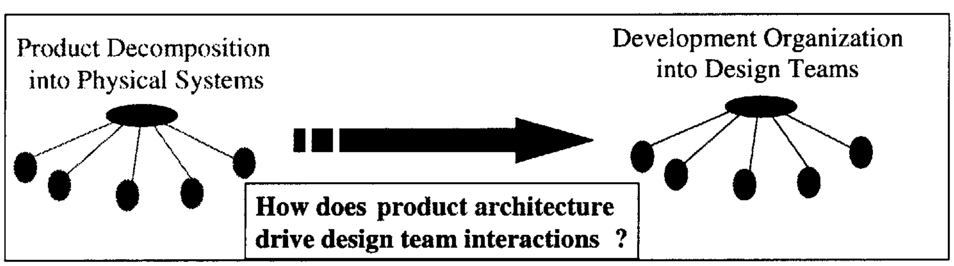

Figure 1-2. Main Research Question

Within this context, we are particularly interested in answering the following questions:

- How accurately can we predict coordination-type communication by analyzing the coupling of product architecture and the structure of the development organization?
- Why do some design interfaces between components not correspond to technical interactions between the teams that design them?
- Why do design teams that develop independent components still engage in technical interaction?
- Is there any difference in the communication patterns of design teams that develop modular systems versus design teams that develop integrative systems?
- How can managers mitigate the negative effects of geographically distributed development teams?

1.4. Thesis Outline

Chapter 2 describes our research method in the context of the development of a large commercial aircraft engine. In chapter 3, we formulate the hypotheses that explain the mismatch between the product architecture and the design team interactions. Chapter 4 contains the results of the analysis of categorical data completed to test the hypotheses posed in chapter 3. A log-linear model that addresses the limitations of the categorical data analysis is presented in chapter 5. An empirical study conducted in the telecommunications industry to address the effects of media use on geographically distributed development organizations is presented in chapter 6. Finally, the conclusions of the thesis and future research directions are outlined in chapter 7.

2. *The Development of a Large Commercial Aircraft Engine*

"Y cuando salto de cubierta
y me abandono a la corriente

Nuevas formas crecen
son tan atractivas
quiero descansar de todo ayer

Y voy flotando por el río
voy envuelto en la corriente.

Hombre al agua"
(Gustavo Cerati/Zeta Bosio, 1990)

In this chapter we describe the research method we used to study the coupling of the architecture of a product and the development organization that designs it. We illustrate our approach in the context of the development of a large commercial engine. After describing how to capture the architecture of the engine, we define the concepts of modular systems and integrative systems. We also describe how we capture the structure of the development organization, and its team interactions. Finally, we illustrate how to map the product architecture and the development organization into a single matrix.

2.1. Research Method

This section describes our method of comparing the architecture of a product with the structure of its development organization. Our approach involves three steps:

1) **Capture the product architecture**. By interviewing design experts[1], we document how the product is decomposed into systems, and these systems into components. Then, we ask them to identify the interfaces required for functionality between the components that comprise the product. We represent such design interfaces in a design interface matrix.

---

[1] These people have a deep understanding of the product architecture. They are not necessarily the people who design the product.

2) **Capture the development organization.** We first identify the design teams responsible to develop the product's components. Then, by surveying key members of the design teams we capture the intensity (i.e. criticality and communication frequency) of the interactions between them. We represent such team interactions in a team interaction matrix.
3) **Compare the product architecture and the development organization.** Finally, we compare the design interface matrix with the team interaction matrix to answer the research questions posed in chapter one.

2.2. Research Site

We apply this approach to the design of a large commercial aircraft engine. The engine studied was a derivative engine. That is, it was the third generation in a family of engines, the 112-inch-fan engine, which is an ultra-high-thrust model. It covers the 74,000 to 98,000-pound-thrust class to meet the current requirements for the Boeing 777 twinjet. It was the launch engine for the 777, entering service in 1995. The model studied is the most powerful commercial engine in the world, and its diameter is nearly as wide as the fuselage of a Boeing 737. Figure 2.1 exhibits a cross-section diagram of the engine studied.

Several factors justified the selection of the project to study. First, the project chosen was a complex design that exhibited explicit decomposition of the engine into systems, and these into components. Furthermore, the engine was comprised of both modular and integrative systems. Second, the way the development team was organized around the architecture of the product facilitated the implementation of our approach. Third, the model studied was the most recent engine program to complete design and development, and almost all team members involved in the initial development phase were still accessible. Finally, the engine studied was part of a family of large commercial engines with two new derivatives planned whose development programs had the potential to gain directly from this analysis. For more details about the general organizational structure of the firm and the data collection process refer to Rowles (1999).

As stated before, the engine studied was the third generation in a family of engines. Indeed, its components were (in average) 60% redesigned with respect to the second generation. On the organizational side, about 80% of the organization that developed the engine studied was also involved with the development of the second generation of this engine.

2.3. Capturing the Product Architecture

The engine analyzed was decomposed into eight systems (see Figure 2.1). Each of these systems was further decomposed into five to seven components each (see Table 2.1). Six out of the eight systems (the fan, the low-pressure compressor, the high-pressure compressor, the burner/diffuser, the high-pressure turbine, and the low-pressure turbine) exhibited a modular architecture in which the interfaces between their components were clearly defined among adjacent components (modular systems). On the other hand, the components of the other two systems (the mechanical components, and the externals and controls) were physically distributed throughout the engine exhibiting an integrative architecture. Components, such as the main shaft and the external tubes are examples of these types of distributed components within the integrative systems. In total, the engine was decomposed into 54 components grouped into eight systems, of which six were *modular systems* and three were *integrative systems*. (In the next section we discuss in detail the basis of this categorization.)

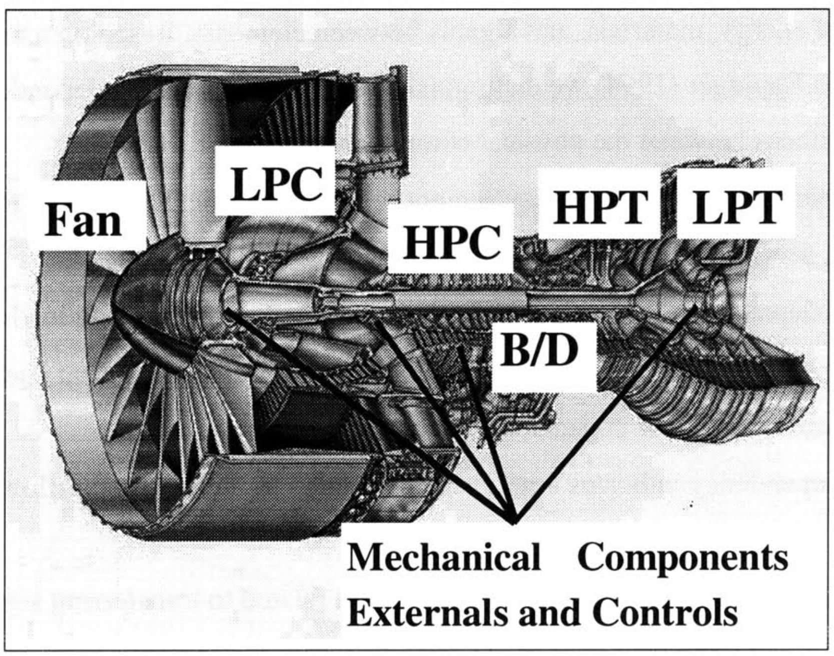

Figure 2-1. Systems of a Large Commercial Aircraft Engine

Table 2-1. Systems and Components of the Engine Studied

System	Number of Components
Fan	7
Low-Pressure Compressor (LPC)	7
High-Pressure Compressor (HPC)	7
Burner and Diffuser (B/D)	5
High-Pressure Turbine (HPT)	5
Low-Pressure Turbine (LPT)	6
Mechanical Components (MC)	7
Externals and Controls (EC)	10

2.3.1. Types of Design Interfaces

After documenting the general decomposition of the product, we proceeded to identify the interfaces between the 54 components of the engine. Researchers in Engineering Design (Suh (1990), Pahl and Beitz (1991)) have modeled functional requirements of product design in terms of exchanges of energy, materials, and signals between elements. Based on a method proposed by Pimmler and Eppinger (1994) we distinguished five types of design dependencies to capture the design interfaces between the physical components:

- **Spatial** dependency indicates a requirement related to physical adjacency for alignment, orientation, servicability, assembly, or weight.
- **Structural** dependency indicates a requirement related to transferring loads, or containment.
- **Energy** dependency indicates a requirement related to transferring heat energy, vibration energy, electric energy, or noise.
- **Material** dependency indicates a requirement related to transferring airflow, oil, fuel, or water.
- **Information** dependency indicates a requirement related to transferring signals or controls.

2.3.2. Criticality of the Design Interface

After design interfaces were identified, we captured the level of criticality for each dependency in the overall functionality of the component in question[2]. Using the five-point scale used by Pimmler and Eppinger (1994) we capture the level of criticality as:

Required (+2): Interface is necessary for functionality.

Desired (+1): Interface is beneficial, but not absolutely necessary for functionality.

Indifferent (0): Interface does not affect functionality.

Undesired (-1): Interface causes negative effects, but does not prevent functionality.

Detrimental (-2): Interface must be prevented to achieve functionality.

2.3.3. Design Interface Matrix

We mapped the design-interface data into a square (54x54) design interface matrix[3]. The identically labeled rows and columns name the 54 components of the engine, and their sequencing follows the physical arrangement of the systems within the engine. Indeed, the systems were sequenced following the airflow through the engine. Each off-diagonal cell of the matrix contains a vector of five values representing the degree of criticality of the five types of design dependency for a single design interface. Hence,

---

[2] This does not force design interfaces between components to be reciprocal.

[3] The design interface matrix can be described as a special form of design structure matrix (DSM). For a formal introduction to DSM refer to Steward (1981) or Eppinger *et al* (1994).

$$A_{54,54} = \textit{Design Interface Matrix}$$

$$a_{i,j} = \begin{bmatrix} c_{i,j}^{spatial} \\ c_{i,j}^{structural} \\ c_{i,j}^{energy} \\ c_{i,j}^{material} \\ c_{i,j}^{information} \end{bmatrix}$$

where, (2.1)

$c_{i,j}^{d}$ = *criticality of the interface of type "d" between components "i" and "j", for functionality of component "i"*

$c_{i,j}^{d} = [-2,-1,0,+1,+2]$

$c_{i,j}^{d}$ *is undefined for* $i = j$

$$A_{54,54}^{b} = \textit{Design Interface Matrix (binary)}$$

$$a_{i,j}^{b} = 1 \quad \text{if } |a_{i,j}| > 0$$

$$a_{i,j}^{b} = 0 \quad \text{if } |a_{i,j}| = 0$$

For graphical simplicity, Figure 2.2 shows a binary version of the design interface matrix. The off-diagonal elements of the matrix are marked with an "X" for each pair of components that shares at least one design interface (any non-zero level of criticality). Reading across a row corresponding to a particular component indicates the other components it depends on for functionality. The diagonal elements are meaningless and are shown to separate the upper and lower triangular portions of the matrix. The boxes around the diagonal indicate the eight system boundaries. Marks inside the boxes represent design interfaces between components of the same system, whereas marks outside the boxes indicate interfaces between components of different systems. The first six systems in the matrix correspond to the six modular systems, while the last two systems correspond to the two integrative systems. Note that the integrative systems have design interfaces with components in every system of the engine.

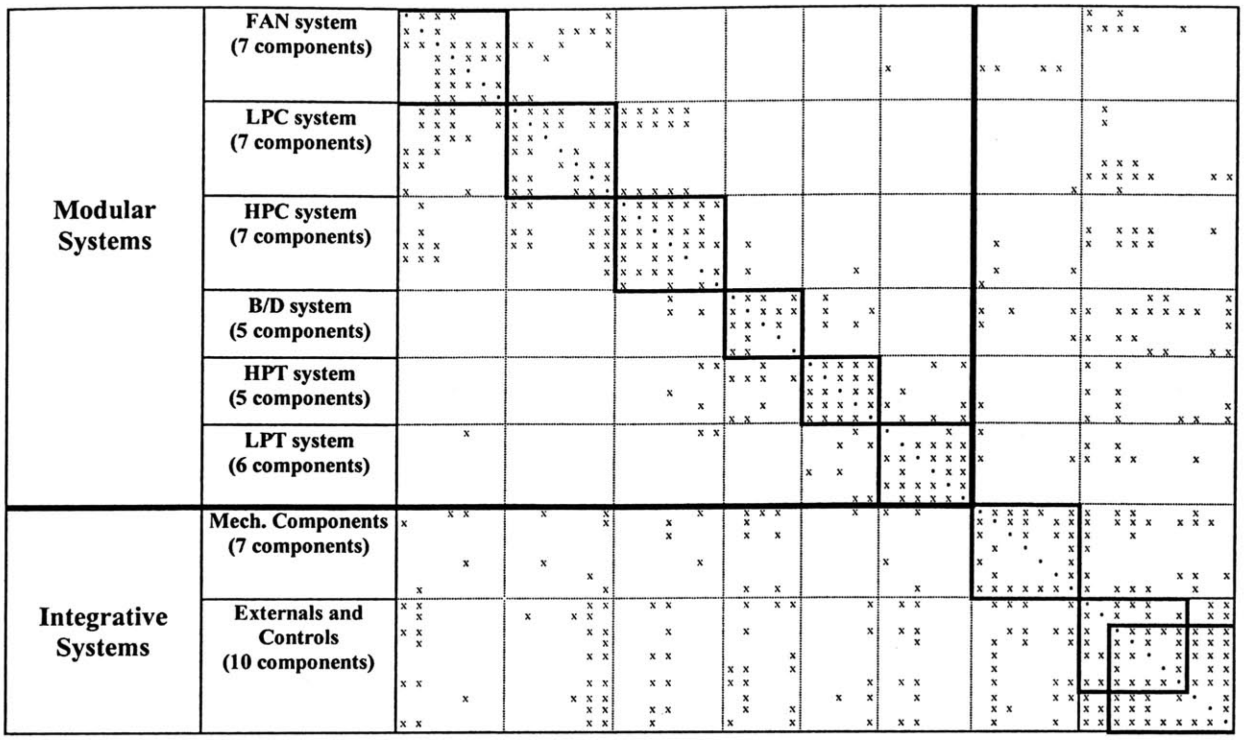

Figure 2-2. Design Interface Matrix (binary)

2.4. Identifying Modular and Integrative Systems

The first six systems of the design interface matrix are those in which design interfaces are primarily among adjacent components (modular systems). In contrast, the mechanical components system, and the externals and controls system exhibit design interfaces distributed throughout the engine (integrative systems).

To determine the degree of modularity (and integrality) of each system we analyzed the distribution of design interfaces across system boundaries. System boundaries are highlighted by the boxes along the diagonal of the design interface matrix. Marks inside the boxes represent design interfaces between components of the same system, whereas marks outside the boxes indicate interfaces between components of different systems. Light boxes throughout the matrix enclose the cross-boundary design interfaces between two systems.

Figure 2.3 shows the number of design interfaces between the externals and controls system and the six modular systems. Similarly, Figure 2.4 shows the number of design interfaces between the mechanical components system and the first six modular systems. To visually compare the difference in distribution between modular and potential integrative systems, Figure 2.5 shows the distribution of the design interfaces between the high-pressure compressor (HPC), which is the least modular of the six systems, and the other modular systems.

Table 2.2 shows the results of the Chi-square test performed to test the alternative hypothesis that "the distribution of design interfaces of the externals and controls system is statistically significantly different from the distribution of the high-pressure compressor system". The test resulted in a χ^2 equal to 29.880 which is greater than the critical value of 9.488 (for α=0.05 and four degrees of freedom). The expected values shown in Table 2.2 are based on the distribution of the design interfaces of the externals and controls system. The actual values are the number of design interfaces of the high-pressure compressor system. Similar results were found when comparing the distribution of cross-system design interfaces of the externals and controls system and the other five modular systems.

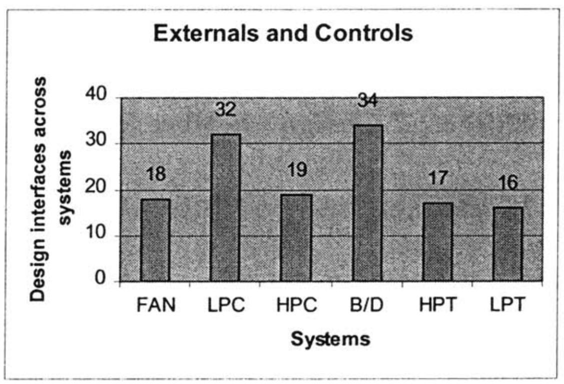

Figure 2-3. Distribution of design interfaces of Externals and Controls system

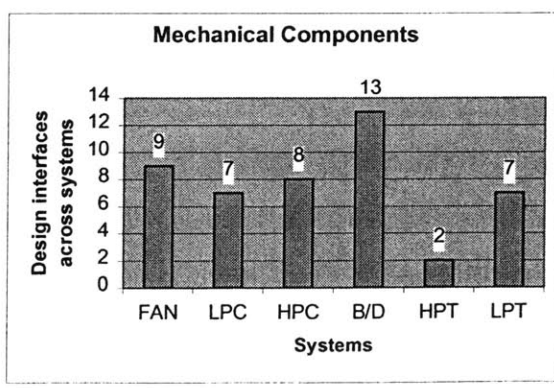

Figure 2-4. Distribution of design interfaces of Mechanical Components system

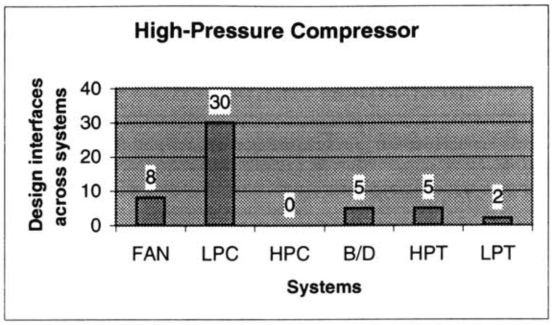

Figure 2-5. Distribution of design interfaces of HPC system

We found that the distribution of design interfaces of the externals and controls system, and the mechanical components system are similarly distributed among the first six modular systems. Table 2.3 shows the results of the Chi-square test performed to test the null hypothesis that "the distribution of design interfaces of the mechanical components system, and the distribution of design interfaces of the externals and controls system are statistically equivalent". The test resulted in χ^2 equal to 6.237 which is smaller than the critical value of 11.070 (for α=0.05 and five degrees of freedom). We cannot, therefore, reject the null hypothesis of no difference in the distribution of design interfaces for the two systems.

Table 2-2. Chi-square test results. Comparing externals and controls system with high-pressure compressor system

System	Expected fraction of design interfaces based on Ext/Controls	Expected number of design interfaces of HPC†	Actual number of design interfaces of HPC	χ^2
FAN	15.38%	7.692	8	0.012
LPC	27.35%	13.675	30	19.488
B/D	29.06%	14.530	5	6.251
HPT	14.53%	7.265	5	0.706
LPT	13.68%	6.838	2	3.423
Total	100.00%	50.000	50	**29.880**

† The fraction of design interfaces between the external and controls system and the fan system is 18 out of a total of 117 design interfaces, that is, 15.38%. Hence, the expected number of design interfaces between the HPC system and the fan system, under the null hypothesis, is the 15.38% of a total of 50 design interfaces, that is, 7.692. The rest of the expected values are determined in similar way.

Table 2-3. Chi-square test results. Comparing externals and controls system with mechanical components system

System	Expected fraction of design interfaces based on Ext/Controls	Expected number of design interfaces of Mech. Comps. †	Actual number of design interfaces of Mech. Comps.	χ^2
FAN	13.24%	6.088	9	1.393
LPC	23.53%	10.824	7	1.351
HPC	13.97%	6.426	8	0.385
B/D	25.00%	11.500	13	0.196
HPT	12.50%	5.750	2	2.446
LPT	11.76%	5.412	7	0.466
Total	100.00%	46.000	46	**6.237**

† The fraction of design interfaces between the external and controls system and the fan system is 18 out of a total of 136 design interfaces, that is, 13.24%. Hence, the expected number of design interfaces between the mechanical components system and the fan system, under the null hypothesis, is the 13.24% of a total of 46 design interfaces, that is, 6.088. The rest of the expected values are determined in similar way.

2.5. Capturing the Development Organization

The organization responsible for the development of the aircraft engine was divided into sixty design teams. Fifty-four of these teams were grouped into eight system-design groups mirroring the architecture of the engine described above. Each of those teams was responsible for developing one of the 54 components of the engine. The remaining six design teams were system integration teams, which had no specific hardware associated with them and whose responsibility was to assure that the engine worked as a whole. Examples of the system integration teams are the rotordynamics team and the secondary flow team.

We capture the system-level integration efforts (both within groups and across groups) of the organization by measuring the intensity of the technical interaction between the design teams involved in the development process. This method is similar to the approach used by McCord and Eppinger (1993). We focused our efforts on capturing coordination-type interactions between design teams. Additionally, the development organization was co-located in a single-floor building, and every team member had access to each other via face-to-face, telephone, and email. Also, the use of a centralized database to save and/or retrieve information was very limited. We explicitly asked respondents not to report knowledge-type or inspiration-type communications.

2.5.1. Team Interaction Intensity

To measure the intensity of each team interaction, we asked at least two key members from each design team to rate the criticality and frequency of their interactions with each of the other teams during the detailed design phase of the engine development project. We used a five-point scale that combines the frequency and criticality of each interaction into a single metric, as shown in Table 2.4. The criticality metric allows asymmetry in the interaction intensity of each pair of design teams. That is, interaction intensity is reported from the respondent's point of view, and we surveyed both parties of each pair to obtain a bilateral view of each interaction.

Table 2-4. Team Interaction Intensity

Criticality\Frequency	**Very Frequently (perhaps daily)**	**Frequently (weekly or biweekly)**	**Infrequently (monthly or less)**	**Never**
Critical: Information cannot be generated by team alone, and its delay or absence causes rework or increases iterations.	5	4	3	0
Important: Information might be generated by team alone with some risk and effort. Its delay or absence negatively affects team performance.	3.5	2	1	0
Routine: Information is important but can be generated by team alone with minimal risk. Its delay or absence has not significant impact in team performance.	2	0.5	0	0

2.5.2. Team Interaction Matrix

We organize the team-interaction data in a square (60x60) team interaction matrix. The identically ordered labels of the rows and columns of this matrix contain the names of each of the design teams. Each cell in the matrix contains the interaction intensity reported by each team. Hence,

$$
\begin{aligned}
&\mathrm{T}_{60,60} = \textit{Team Interaction Matrix} \\
&t_{i,j} = \textit{team interaction intensity } [0,5] \textit{ reported by team "} i \textit{" about its interaction with team "} j \textit{".} \\
&t_{i,j} \textit{ is undefined for } i = j \\
&\mathrm{T}^{b}_{60,60} = \textit{Team Interaction Matrix (binary)} \\
&t^{b}_{i,j} = 1 \;\; \textit{if } t_{i,j} > 0 \\
&t^{b}_{i,j} = 0 \;\; \textit{if } t_{i,j} = 0
\end{aligned}
\tag{2.2}
$$

Figure 2.6 shows a binary team interaction matrix with off-diagonal cells marked "O" to indicate each non-zero team interaction revealed. Reading across a particular row indicates with which other teams the surveyed team interacted.

The 60 design teams are organized into groups that mirror the product architecture structure. As shown in Figure 2.6, associated with the six modular systems are corresponding groups of design teams. Similarly, the two integrative systems have their two corresponding groups of design teams. Finally, there are six system integration teams that are not responsible for designing any specific engine's component but they are in charge of integrating all the components into a whole. The boxes around the diagonal indicate the organizational boundaries between the eight design groups. Marks inside the boxes indicate within-boundaries team interactions, which we associate to within-group system-level integration effort. On the other hand, marks outside the boxes indicate cross-boundaries team interactions, which we associate to across-group system-level integration effort.

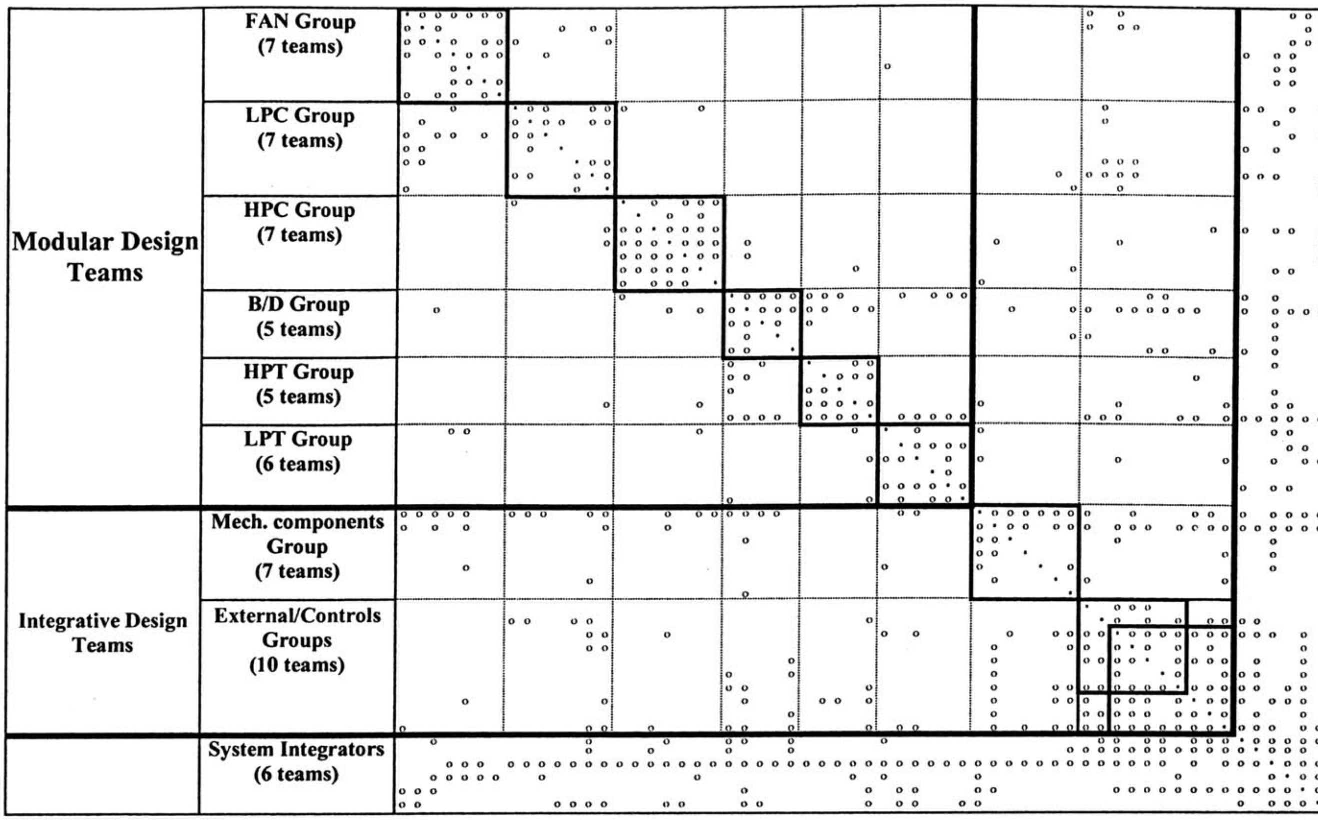

Figure 2-6. Team Interaction Matrix (binary)

2.6. Comparing Product Architecture and Development Organization

The one-to-one assignment of the 54 components to the 54 design teams allows the direct comparison of the design interface matrix with the team interaction matrix. In Appendix C, we present an algebraic model that allows us to perform this comparison for the general case when the assignment is not one-to-one. Figure 2.7 shows how, by overlapping the design interface matrix over the team interaction matrix, we obtain the resultant matrix. The resultant matrix is exhibited in Figure 2.8.

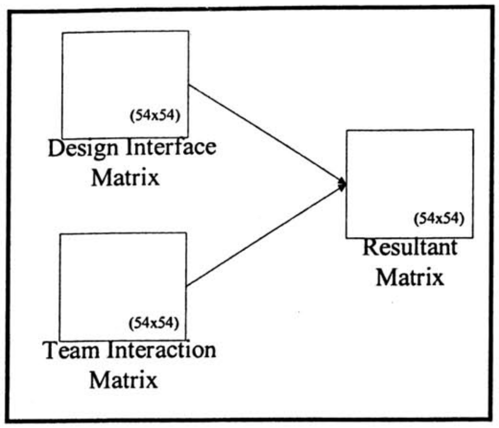

Figure 2-7. Comparing Product Architecture and Development Organization Interactions

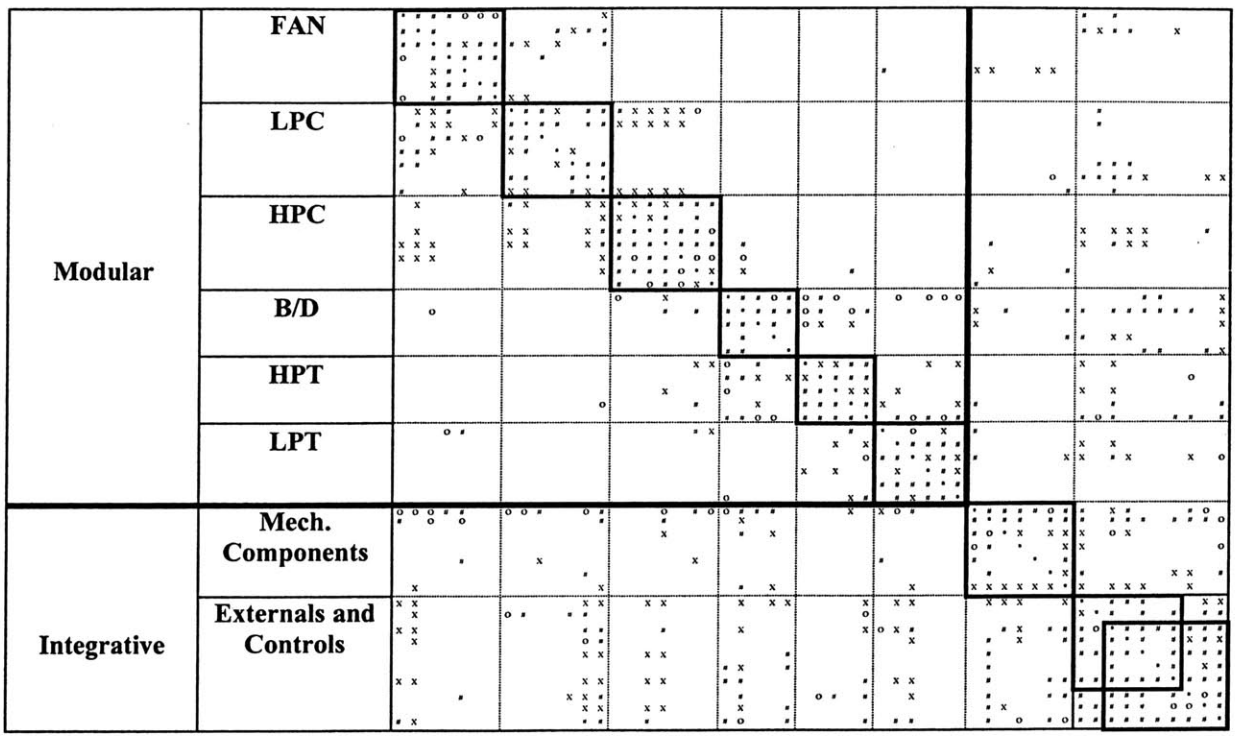

Figure 2-8. Resultant Matrix

Legend for Figure 2.8:

X: Design interface with no team interaction
O: Team interaction with no design interface
#: Both design interface and team interaction
(Blanks): No design interface and no team interaction
***:** Diagonal elements (meaningless)

3. Formulating the Hypotheses

"Pero yo como soy tan sencillo
Pongo en claro esta trovada
Compay, yo no dejo el trillo
Para meterme en cañada."
(en De Camino a La Vereda (1950s), de Ibrahim Ferrer)

The resultant matrix, exhibited in Figure 2.8, provides the basis for the analysis completed to answer our research questions. Figure 3.1 exhibits the four possible outcomes for each cell of the resultant matrix. Two positions in the 2x2 matrix shown in Figure 3.1 represent the expected cases in which either design interfaces are matched by team interactions ("#" cell), or absence of team interaction corresponds to lack of design interface (blank cell). However, the two unexpected cases ("X" and "O" cells) are far more interesting. In the "X" cell we find the cases in which design interfaces are not matched by team interactions. In the "O" cell we find the cases in which team interactions were not predicted by design interfaces.

Team Interaction			
	NO	X	
	YES	#	O
		YES	NO
		Design Interface	

Figure 3-1. Four possible values of each cell of the resultant matrix

While we expect the majority of the cells of the resultant matrix to contain "blank" and "#" cells, we will focus our analysis on the unexpected cases. In this section we present several hypotheses which can explain occurrence of the two types of unexpected cases described above. First, we hypothesize possible explanations for the cases when design interfaces are not matched by team interactions. Then, we hypothesize possible reasons that may explain the cases when team interactions are not predicted by design interfaces.

3.1. Design Interfaces That Are Not Matched by Team Interactions (The "X" Cells)

We hypothesize that whether or not the known design interfaces are matched by team interactions (the YES column in Figure 3.1) is contingent upon the following effects:

- Effect due to organizational boundaries
- Effect due to design interface strength
- Effect due to design interface type
- Effect due changes with respect to previous generation
- Effect due to indirect team interactions
- Effect due to cross-membership of team members
- Effect due to design escape
- Measurement errors

3.1.1. Effect due to organizational boundaries

Organizational boundaries are defined by the way design teams are grouped into system teams. These boundaries impose communication barriers that prevent design teams from interacting (Allen (1977), Van den Bulte and Moenaert (1998), Sosa *et al* (1999)). Empirical evidence from R&D organizations suggests that interactions within organizational boundaries are more likely to occur than across organizational boundaries. People within such boundaries are subjected to organizational bonds that promote the development of a language and an identity inherent to the group. Indeed, in chapter seven we provide empirical evidence from the telecommunications industry that higher communication frequency is found in pairs that share organizational bonds. Allen (1977) found higher probability of engineers (in R&D organizations) engaging in technical communication when they share organizational bonds. Thus, we expect the following hypothesis to hold true:

H3.1: Design interfaces within organizational boundaries are more likely to be matched by team interactions than design interfaces across organizational boundaries.

3.1.2. Effect due to design interface strength

Recent research suggests that a greater degree of design interdependence results in more communication. Allen (1997) claims that the degree of interdependence between engineers' work is directly related to the probability that they engage in frequent technical communication. At the task level, Smith and Eppinger (1997) use the strength of task interdependency to identify the set of activities that require higher effort to coordinate. Loch and Terwiesch (1998) use an analytical approach to suggest that communication frequency increases with the level of dependence. These results are consistent with the empirical evidence presented by Adler (1995) and the numerical approach presented by Ha and Porteus (1995). In the empirical study presented in chapter seven, we also show that communication frequency increases with the degree of interdependence, independently of the communication medium used. Thus far, we expect to find empirical support for the following hypothesis:

H3.2: Strong (i.e. critical and multi-dependency) design interfaces are more likely to be matched by team interactions than weak (i.e. non-critical and few dependencies) design interfaces.

3.1.3. Effect due to design interface type

According to the type of design dependency, we classify design interfaces into two major categories:

Spatial-type design interfaces, which involve spatial dependencies only.

Transfer-type design interfaces, which involve structural and/or energy and/or material dependencies. Information dependencies are not included because they are not present in modular systems.

Henderson and Clark (1990) refer to *communication filters* as the mechanism for screening the most crucial information. Adapting this concept to our context, we expect some design teams to handle a larger proportion of some types of design interfaces than other ones. Hence, we want to investigate whether there is a difference in the way modular and integrative design teams handle these two types of design interfaces. Specifically, we would like to test the following hypothesis:

H3.3: Spatial-type design interfaces are more important for modular systems design than for integrative system design, whereas transfer-type design interfaces are more important for integrative system design than for modular system design.

3.1.4. Effect due to changes with respect to previous generation

Uncertainty, defined by Galbraith (1973) as "the difference between the amount of information required to perform the task and the amount of information already possessed", leads directly to information flow through the communication network. An important source of uncertainty when designing a derivative product is due to changes with respect to the previous generation. Changes may be attributed to two sources. First, changes in the product, and second, changes in the organization.

Loch and Terwiesch (1998) propose an analytical approach which suggests that average communication level increases with uncertainty and dependence. They refer to uncertainty as the rate of modifications of the upstream task (similar to *upstream evolution* in Krishnan *et al* (1997)), while dependence refers to the impact caused on downstream task due to changes in the upstream activity (similar to downstream sensitivity in Krishnan *et al.* (1997)). We define percentage of redesign of the interface as the portion of the interface that differs from the previous model of the product. Following previous work (Krishnan *et al.* (1997), Loch and Terwiesch (1998), Carrascosa *et al.* (1999)) on the categorization of information transfer between concurrent engineering activities, the redesign of the interface (i,j) can be defined in terms of the *change of the design interface* caused by changes of component j, and the *impact* of those changes on component i. Hence, we expect the following hypothesis to hold true:

H3.4a: The larger the percent of interface redesign the higher the probability of such a design interface to be matched by team interactions.

Another source of uncertainty between two teams handling a particular design interface refers to the involvement of the teams in previous design of the product. Even if the design interface is not redesigned at all, teams that have not been exposed to previous designs of the product still need to interact to solve their interfaces. Hence, we hypothesize the following:

H3.4b: The lower the involvement of design teams in previous design of the product, the higher the probability that their team interactions would match their design interfaces.

3.1.5. Effect due to indirect team interactions

Two design teams may not directly communicate to implement their design interfaces because they instead communicate indirectly about their interface through a third team. Indirect communication may also occur when design teams use a centralized database in which information is stored and retrieved by teams in the organization.

In our study, we define potential indirect team interactions as those that might occur between design teams that do not interact directly with each other, but have a common interaction with another team. For example, team B has a potential indirect interaction with team A because both of these teams interact with team C (see Figure 3.2).

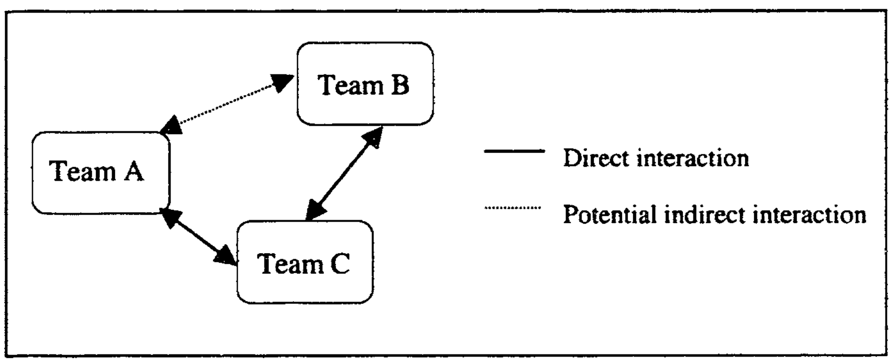

Figure 3-2. Potential Indirect Team Interaction between Teams A and B

It therefore follows that the number of potential indirect team interactions for a given pair of teams is proportional to the probability that technical information is transmitted through some of those indirect interactions. We should expect empirical support for the following hypothesis:

H3.5: Teams that do not interact directly have a greater likelihood of sharing a design interface if they have a greater number of potential indirect team interactions.

3.1.6. Effect due to cross membership of team members

We define cross membership when team members participate in the design of various components simultaneously. If two design teams share the membership of one or more team members we should expect a lower interaction intensity between those teams because part of the information transfer would be carried out through the shared team members.

In our study, this effect was not particularly relevant since most of the team members were assigned to only one design team.

3.1.7. Effect due to design escape

We define design escape when design interfaces were known but not addressed during the design period. We do not have quantitative data to address what portion of the design interfaces not matched by team interactions were actually design escapes, but it is certainly a reason that might explain some of those cases.

3.1.8. Effect due to measurement errors

Capturing the design interfaces and the interactions between the teams was a very difficult task where the potential of measurement errors was always present. We recognize that not all coordination-type interactions were reported during the survey and some design interfaces might have been mistakenly documented.

3.2. Team Interactions That Are Not Predicted by Design Interfaces (The "O" Cells)

We now turn our attention to the team interactions (the YES row in Figure 3.1) and explore why some of these interactions are <u>not</u> predicted by design interfaces. We hypothesize three contingencies to explain the unpredicted interactions:

- Effect due to system boundaries
- Effect due to secondary design interfaces
- Effect due to measurement errors

3.2.1. Effect due to system boundaries

System boundaries are defined by the way components form systems. Such boundaries may impose architectural knowledge barriers which inhibit explicit identification of cross-system design interfaces by the design experts. Nevertheless, in order to develop working systems, the teams learn of their needs to interact and do so. This results in team interactions that are not predicted by the design interfaces. Hence, we should expect a higher percentage of unknown design interfaces across system boundaries. We formulate the following hypothesis:

H3.6: Unknown design interfaces are more likely to occur across system boundaries, hence team interactions across system boundaries are less likely to be predicted by design interfaces than team interactions within system boundaries.

3.2.2. Effect due to secondary design dependencies

We define secondary design dependencies as those which occur between components that do not depend on each other directly, but might depend on each other through other components. For example, component A does not depend directly on component B, but it might share a secondary design dependency with component B because component C depends on component B and component A depends on component C (see Figure 3.3).

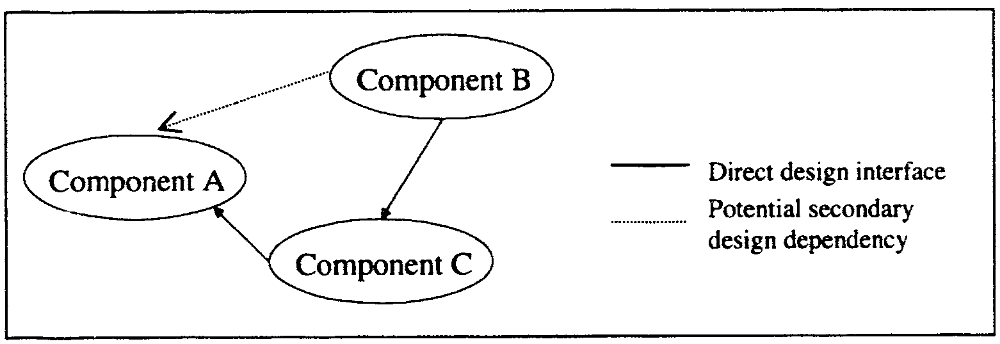

Figure 3-3. Potential Secondary Design Dependency

By considering potential secondary design dependencies we recognize that system-level dependencies are important, and that some would be handled by the design teams. If the chances of finding system-level dependencies (between components that do not share a direct design interface) are proportional to the number of secondary design dependencies, we should expect the following hypothesis to be true:

H3.7: The higher the number of potential secondary design dependencies between components with no direct design interface, the higher the probability that the design teams that develop those components interact.

3.2.3. Effect due to measurement errors

Some of the unexpected cases can be attributed to errors occurred during the data collections. Specifically, we understand that not all known design interfaces were captured and some team interactions might have been erroneously reported.

4. Categorical Data Analysis

"Things should be made as simple as possible, but no simpler"
Albert Einstein

Categorical data analysis takes place when data comes in the form of counts in various categories. The data are usually displayed in cross-classified tables of counts, commonly referred as contingency tables. Many statistics textbooks (Daniel and Terrel (1995), Rice (1995)) focus on the analysis of such data in the special case of two-way cross-classifications. However, when we look at several categorical variables simultaneously, we deal with multidimensional contingency tables, with each variable corresponding to one dimension of the table (Bishop *et al.* (1975), Fienberg (1980)).

In the next section we provide an overview of the types of statistical tests completed in the context of categorical data analysis. Then, we describe the Bernoulli probability distribution as the basis of the fundamental assumption underlined in the analysis presented in this chapter. Finally, we present the results of the statistical analysis completed to test the hypotheses posed in the previous chapter.

4.1. Statistical Tests

We conduct three types of tests with various samples of our data. In general, we carry out a test of independence to test some of the hypotheses posed in chapter 3. Then, we complete a test of homogeneity to test the moderating effects due to the modularity and integrative nature of the systems. Finally, we test the moderating effects due to organizational and system boundaries by using log-linear models for three-dimensional contingency tables.

4.1.1. The Chi-Square Test of Independence

We perform this type of test when evaluating the null hypothesis that two criteria of classification, when applied to a set of data points, are *independent*. If two criteria of classification are not independent, there is an *association* between them. We will construct two-dimensional contingency tables containing a sample of n data points cross-classified according to two criteria specified by the hypothesis to test. The first criterion is based on the selection of the sample (for example, whether a design interface is matched or not by a team interaction). The second criterion is defined by the hypothesized effect (for example, a design interface is either within or across organizational boundaries). Hence, the data support our hypothesis when the two criteria of classification are *statistically independent* of each other, and their *association* is in line with the hypothesized effect.

A two-dimensional contingency table contains I rows and J columns. The observed value corresponding to cell (i,j) of the table is denoted by x_{ij}, while the expected value corresponding to the same cell is denoted by $\hat{m}_{ij}$. Hence, the Pearson's chi-square statistic can be determined as:

$$X^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(x_{ij} - \hat{m}_{ij})^2}{\hat{m}_{ij}} \qquad (4.1)$$

To determine the expected values of the contingency table under the hypothesis of independence we use the principles of probability. It can be shown that if two criteria of classification are independent, a joint probability is equal to the product of the two corresponding marginal probabilities (Daniel and Terrel, 1995). That is,

$$\hat{m}_{ij} = (\frac{x_{i+}}{n})(\frac{x_{+j}}{n})n \qquad (4.2)$$

Our notation for marginal totals follows Fienberg (1980) notation so that when we add over a variable we replace the corresponding subscripts by a "+".

(x_{i+}/n) = marginal probability that randomly picked data point ij is characterized by the criterion corresponding to row i

(x_{+j}/n) = marginal probability that randomly picked data point ij is characterized by the criterion corresponding to column j

To determine the number of degrees of freedom, we note that there are IJ-1 independent cells in the table because the grand total (n) is fixed, and (I-1) + (J-1) independent parameters are estimated from the data. Since we know that

df = number of independent cells - number of independent parameters estimated

Then,

$df = (I\text{-}1)(J\text{-}1)$ (4.3)

4.1.2. The Chi-Square Test of Homogeneity

This type of test is used to explore the hypothesis that several populations are homogenous with respect to some characteristic. We used this test to explore the difference between modular and integrative systems. By splitting the samples used in the test of independence into two subgroups, one for modular systems and another for integrative systems, we investigate whether there is any statistically significant difference between modular and integrative systems with respect to the hypothesized effect.

The fundamental difference between the test of independence and the test of homogeneity is that, in the former we use one sample, while in the latter we use two or more samples. To get the expected frequencies in the test of homogeneity we pool the values from the sample data (Daniel and Terrel, 1995).

4.1.3. A Loglinear Model for Three-Dimensional Contingency Tables

In order to explore the interaction effects due to organizational and system boundaries and the other hypothesized effects we construct three-dimensional contingency tables whose expected values are given by log-linear models that include the mentioned interaction effects.

If we assume that the three variables on the contingency tables are independent then, by analogy with the model of independence in two dimensions we can obtain the estimates of the expected count for the (i,j,k) cell as follows:

$$\cdot_{ij} = \left(\frac{x_{i++}}{n}\right)\left(\frac{x_{+j+}}{n}\right)\left(\frac{x_{++k}}{n}\right)n \qquad (4.4)$$

Taking the natural logarithms yields to

$$ln\ \hat{m}_{ijk} = ln\ x_{i++} + ln\ x_{+j+} + ln\ x_{++k} - 2 \cdot ln\ n \qquad (4.5)$$

Equation (4.5) can be re-written in an ANOVA-type notation as follows:

$$ln\ \hat{m}_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} \qquad (4.6)$$

where,

$$\sum_{i=1}^{I} u_{1(i)} = \sum_{j=1}^{J} u_{2(j)} = \sum_{k=1}^{K} u_{3(k)} = 0 \qquad (4.7)$$

If the three variables are not independent, other models including the interaction effects between the variables can be formulated. We are particularly interested in the interaction effects of organizational and system boundaries with the hypothesized effect. Hence, if *j* denotes the hypothesized effect and *k* denotes the effect of organizational and system boundaries the log linear model can be specified as follows:

$$\ln \hat{m}_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)} \qquad (4.8)$$

The model specified by 4.8 consider that variables 2 and 3, taken jointly, are independent of variable 1 (Finberg, 1980). In our context, variable 2 captures the hypothesized effect with respect to variable 1 while variable 3 captures the effect due to organizational and system boundaries. For example, when analyzing the joint effects of design interface strength and organizational boundaries on the probability of a design interface being matched by team interaction, variable 1 indicates whether a design interface is matched or not by team interaction, variable 2 indicates whether such design interface is weak or strong, and variable 3 indicates whether such design interface is either within or across organizational boundaries.

4.2. The Bernoulli Probability Distribution

In order to carry out the statistical tests described in the previous section, we need to assume that the cells in both the design interface matrix and the team interaction matrix are *statistically independent*. In this section we provide the theoretical background that supports that assumption (for details refer to Wasserman and Faust, 1994).

Let us define $\boldsymbol{G_d}$ a particular directed graph with g nodes, and $\boldsymbol{X}$ its adjacency matrix. The set of all possible digraphs with g nodes will be denoted by $\boldsymbol{G_d(N)}$. The digraphs that we consider in this chapter are assumed to be random, so that $\boldsymbol{X}$ itself is a random matrix. Since we concentrate our analysis in binary relations (i.e. the binary design interface matrix and the binary team interaction matrix), each cell of the **X** is either 0 or 1. Since there are g elements, each of whom may have ties to the g-1 other elements, there are $g(g-1)$ possible non-zero cells in $\boldsymbol{X}$. Hence, there must be $2^{g(g-1)}$ different labeled matrices. Clearly, the number of possible realizations of the random matrix $\boldsymbol{X}$ is very large, even for small g (for example for g=10, there are $2^{90}=1.23 \times 10^{27}$ possibilities).

The simplest distribution on $G_d(N)$ is the uniform distribution, in which every realization is equally likely. Hence, the uniform probability function is

$$P(\mathbf{X} = \mathbf{x}) = \frac{1}{2^{g(g-1)}} \qquad (4.9)$$

That is, the probability that matrix X with g elements equals a specific "configuration" of choices ($\mathbf{x}$) is $1/2^{g(g-1)}$. Hence, each of the elements of the sample space has an equal probability of occurring. Under this distribution the cells of the matrix can be described as *statistically independent* Bernoulli random variables with probabilities of choices all equal to 1/2:

$$P(X_{ij}=1) = 1/2, \text{ if } i \neq j$$
$$P(X_{ij}=1) = 0, \text{ if } i = j \qquad (4.10)$$

That is, all elements of the matrix X are independent of all other elements, and the probability distribution of any one of the elements is the simplest possible distribution (the Bernoulli distribution with equal probabilities).

The uniform distribution can be generalized to a family of Bernoulli distributions by varying the probability that any cell of the random matrix X equals to one. Thus, the uniform distribution discussed above is a special case of the Bernoulli distribution. In the Bernoulli distribution, the cells of X are assumed to be Bernoulli random variables with probabilities

$$P(X_{ij}=1) = P_{ij}, \text{ if } i \neq j$$
$$P(X_{ij}=1) = 0, \quad \text{if } i = j \qquad (4.11)$$

where $0 \leq P_{ij} \leq 1$. The $\{ P_{ij} \}$ may differ from element to element to allow some nodes to choose other nodes with different probabilistic tendencies. If the $\{ P_{ij} \}$ are all equal, but not equal to 1/2, the distribution is not uniform.

If we assume that the random variables representing the $g(g-1)$ possible arcs in the digraph follow a Bernoulli distribution, we can test particular hypothesis about L (the number of arcs in the digraph, that is, the number of non-zero cells in matrix X).

Let us assume that X follows a Bernoulli distribution with a constant probability P_o, and L, the number of arcs or non-zero cells of X, is a random variable, with a binomial distribution with parameters $g(g-1)$ and P_o. It follows that

$$E(L) = P_o\, g(g-1) \qquad (4.12)$$

and the variance of L is

$Var(L) = P_o(1 - P_o)g(g\text{-}1)$ (4.13)

Thus, we want to test the hypothesis

H_o: $L \sim Bin(g(g\text{-}1), P_o)$

Assuming that g is large enough to support the large sample theory for the binomial distribution, the test statistic for this hypothesis can be determined as follows (Wasserman and Faust, 1994):

$$z_l = \frac{l - P_o \cdot g(g-1)}{\sqrt{P_o(1-P_o) \cdot g(g-1)}} \qquad (4.14)$$

z_l is approximately standard normal with a mean of 0 and a variance and standard deviation of 1. The p-value for the significance test of hypothesis H_o can be found by determining the probability that a standard normal variable exceeds the value of z_l calculated according to 4.12. Since this is a two-tailed significance test, the p-value for the hypothesis is twice this probability.

4.2.1. Estimating the Constant Probability Governing a Bernoulli Distribution

In order to strengthen our assumption that both the design interface matrix and the team interaction matrix follow a Bernoulli distribution, we estimate the constant probability governing the presence/absence of a non-zero cells on each of these matrices. By substituting P_o for the unknown constant probability P on equations 4.12 and 4.13 we can express the expected values of L and its variance as a function of P. The maximum likelihood estimate of this unknown probability is simply the empirical fraction of non-zero cells present in the corresponding matrix, that is $l/(g(g\text{-}1))$. With this estimate we can also estimate E(L) and $Var(L)$. We can also calculate a confidence interval for the unknown value of P as follows:

$$\hat{P}_{lower} \leq P \leq \hat{P}_{upper}$$

where,

$$\hat{P}_{lower} = \hat{P} - z_{\alpha/2}\sqrt{\hat{P}(1-\hat{P})/g(g-1)} \qquad (4.15)$$

and

$$\hat{P}_{upper} = \hat{P} + z_{\alpha/2}\sqrt{\hat{P}(1-\hat{P})/g(g-1)} \qquad (4.16)$$

where z_α is the upper α x 100 percentage point of the standard normal distribution (Wasserman and Faust, 1994).

Specifically, for the design interface matrix we estimate the probability of a non-zero cell as:

$$\hat{P}_{design\,interfaces} = \frac{569}{54(53)} = 0.199 \qquad (4.17)$$

The endpoints for a 95 percent confidence interval (using $z_{0.025}$=1.96) for the unknown P for the design interface matrix are:

$\hat{P}_{lower} = 0.184$ and $\hat{P}_{upper} = 0.213$

Similarly, for the team interaction matrix we estimate the probability of a non-zero cell as

$$\hat{P}_{teaminteractions} = \frac{423}{54(53)} = 0.148 \qquad (4.18)$$

with a 95 confidence interval specified by the following endpoints:

$\hat{P}_{lower} = 0.135$ and $\hat{P}_{upper} = 0.161$

The Bernoulli distribution is a very simple distribution that does not take into account tendencies of differential attraction of the nodes nor reciprocation effects. Even though both the design interface matrix and the team interaction matrix seem to exhibit those effects, we are going to postpone to chapter five the discussion on how to incorporate such effects in our statistical analysis. For the analysis completed in this chapter, we will assume that the design interface matrix and the team interaction matrix follow a Bernoulli distribution with estimated constant probabilities given by (4.17) and (4.18), respectively.

4.3. Overall Results

Figure 4.1 summarizes the binary results shown in the resultant matrix (Figure 2.8). As expected, the majority of the cases (90% of the cells) are the cases when known design interfaces were matched by team interactions (349 "#" cells), and the cases with no design interfaces and no reported team interactions (2219 blank cells). The unexpected cases accounted for 10% of the cells; those were the cases when known design interfaces were not matched by team interactions (8%, or 220 "X" cells), and the cases when reported teams interactions were not predicted by design interfaces (2%, or 74 "O" cells).

Team Interactions	NO (2439)	**X** (220)	(2219)
	YES (423)	**#** (349)	**O** (74)
		YES (569)	NO (2293)
		Design Interfaces	

Figure 4-1. Overall Results

Among the 569 design interfaces, we found that 61% of those interfaces were matched by team interactions. Among the 423 team interactions, we found that 83% of those team interactions were predicted by design interfaces. Additionally, of the 2293 cases in which no design interfaces were known, 97% did not report team interactions. Finally, of the 2439 cases in which no direct team interactions were reported, 91% did not correspond to design interfaces.

4.4. Testing the Nominal Hypothesis

Before testing the hypotheses posed in chapter 3, we first test the nominal null hypothesis that "*a team interaction is independent of whether there is a design interface associated to it*". We first assume that both design interface matrix and the team interaction matrix follow Bernoulli probability distributions with corresponding constant probability estimates given by 4.17 and 4.18. The joint probability distribution that predicts the cases where a design interface is matched by a team interaction yields to 3% of the 2862 cases have a chance of being "#" cell. Figure 4.2a shows the expected values under the null hypothesis previously stated, and Figure 4.2b exhibits the $\chi 2$ resulted when comparing the expected with the actual values of Figure 4.1. As expected, the $\chi 2$ obtained are remarkably greater than the critical value of 6.635 (for one degree of freedom and $\alpha = 0.01$), therefore we strongly reject the nominal null hypothesis stated above.

Expected values

Team Interactions	Design Interfaces: YES	Design Interfaces: NO
NO	484.9	1954.1
YES	84.1	338.9

$\chi 2$ values

Team Interactions	Design Interfaces: YES	Design Interfaces: NO
NO	144.7	35.9
YES	834.4	207.1

Figure 4-2. a) Expected values under H_o. b) χ^2 values.

Although we provide strong evidence to support the hypothesis that the existence (or lack) of a design interface is matched by the presence (or absence) of a team interaction, we focus on understanding the unexpected mismatches between design interfaces and team interactions.

In order to test the hypotheses posed in the previous section we perform analyses of frequencies on various subsets of the data, relevant to each hypothesized effect (see Figure 4.3).

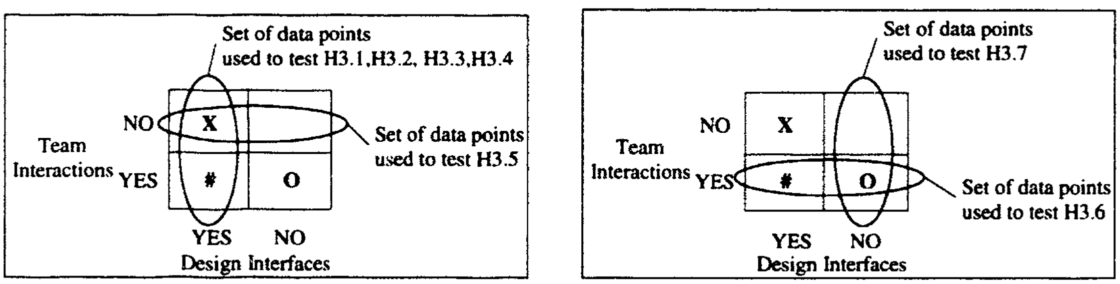

Figure 4.3. a) Data sets used to test H3.1, H3.2, H3.3, H3.4, and H3.5. b) Data sets used to test H3.6, and H3.7.

In the next sub-section, we test the hypotheses that explain the cases where known design interfaces were not matched by team interactions (H3.1, H3.2, H3.3, H3.4, and H3.5). Then, we show the results of testing the hypotheses that help us to explain the cases where reported team interactions were not predicted by design interfaces (H3.6, and H3.7).

4.5. Design Interfaces Not Matched by Team Interactions

The results presented in this section help us to understand the 220 cases ("X" cells) in which known design interfaces were not matched by team interactions.

4.5.1. Effects due to organizational boundaries

As stated in H3.1, we hypothesize that organizational boundaries between design teams may increase the likelihood that teams which share design interfaces across such boundaries will fail to execute coordination-type communications. For the project studied, organizational boundaries are created by the way design teams were assigned to system groups such as the fan group or low-pressure compressor group. Since the organizational structure mirrors the architecture of the product, organizational boundaries and system boundaries are equivalent and they are identically highlighted in each of the matrices by the large, square boxes along the diagonals.

In order to test hypothesis H3.1, we categorize the 569 design interfaces (YES column of Figure 4.1) according to the following two criteria:

- **First criterion**: Whether a design interface is matched by a team interaction or not.
- **Second criterion**: Whether a design interface is either within or across organizational boundaries.

We display the cross-classification of the data in a contingency table (Table 4.1) to perform a chi-square test of independence. The results shown in Table 4.1 allow us to reject the null hypothesis that the two criteria mentioned above are independent. The test resulted in a $\chi 2$ of 63.101, well above the critical value of 6.635 (for one degree of freedom and $\alpha = 0.01$). As shown in Table 4.1, of the 231 design interfaces within organizational boundaries, 81% of them were matched by team interactions whereas of the 338 design interfaces across organizational boundaries, only 48% were matched by team interactions. Therefore the empirical evidence supports hypothesis H1 that design interfaces within organizational boundaries are more likely to be matched by team interactions than when interfaces are across organizational boundaries.

Table 4-1. Chi-square test of independence. Effects of organizational boundaries

	Total	Expected number (fraction) of design interfaces matched by team interactions	Expected number (fraction) of design interfaces not matched by team interactions	Actual number (fraction) of design interfaces matched by team interactions	Actual number (fraction) of design interfaces not matched by team interactions	χ^2 of design interfaces matched by team interactions	χ^2 of design interfaces not matched by team interactions
Design interfaces within organizational boundaries	231	141.685 (61.34%)	89.315 (38.66%)	187 (80.95%)	44 (19.05%)	14.493	22.991
Design interfaces across Organizational boundaries	338	207.315 (61.34%)	130.685 (38.66%)	162 (47.93%)	176 (52.07%)	9.905	15.713
Total	569	349.000	220.000	349	220	**24.398**	**38.703**

H_o: Design interfaces within organizational boundaries are as likely to be matched by team interactions as design interfaces across organizational boundaries.

χ^2 = **63.101** Critical $\chi^2_{(0.99,1)} = 6.635$ Since $\chi^2 >$ Critical $\chi^2_{(0.99,1)}$, we reject H_o.

4.5.1.1. Moderating effects due to systems modularity

Given the distributed nature of integrative systems we expect integrative design teams to be less affected by organizational boundaries. Since the design interfaces of integrative systems are distributed throughout the engine, we anticipate integrative design teams to be more accustomed to cross organizational boundaries than do modular design teams. Indeed, we hypothesize that integrative design teams handle a larger portion of design interfaces across organizational boundaries than do modular design teams.

The chi-square tests of homogeneity shown in Table 4.2 resulted in a χ^2 equal to 0.095 for the cases within boundaries, which is well below the critical value of 6.635 (for α=0.01 and one degree of freedom). On the other hand, for the cases across organizational boundaries χ^2 equaled 8.740, which is well above the critical value.

These results did not allow us to reject the null hypothesis that design interfaces within organizational boundaries are equally handled by teams that design modular systems and by teams that design integrative systems. However, analyzing the cases across organizational boundaries we found that integrative design teams handle a statistically significant higher portion of design interfaces than modular design teams do. Specifically, integrative design teams matched 53.51% of the cross-system design interfaces while modular design teams matched 36.36% of their cross-system design interfaces.

Table 4-2. Chi-square test of homogeneity. Effects of organizational boundaries †

	Total	Expected cases of design interfaces matched by team interactions	Expected cases of design interfaces not matched by team interactions	Actual cases of design interfaces matched by team interactions	Actual cases of design interfaces not matched by team interactions	$\chi 2$ of cases of design interfaces matched by team interactions	$\chi 2$ of cases of design interfaces not matched by team interactions
Design interfaces within organizational boundaries (Modular systems)	137	110.905 (80.95%)	26.095 (19.05%)	110 (80.29%)	27 (19.71%)	0.007	0.031
Design interfaces within organizational boundaries (Integrative systems)	94	76.095 (80.95%)	17.905 (19.05%)	77 (81.91%)	17 (18.09%)	0.011	0 046
Total	231	187.000	44.000	187	44	**0.018**	**0.077**
Design interfaces across organizational boundaries (Modular systems)	110	52.722 (47.93%)	57.278 (52.07%)	40 (36 36%)	70 (63.64%)	3.070	2.826
Design interfaces across organizational boundaries (Integrative systems)	228	109.278 (47.93%)	118.722 (52.07%)	122 (53.51%)	106 (46.49%)	1.481	1 363
Total	338	162.000	176.000	162	176	**4.551**	**4.189**

† Expected values are determined with the pooled data which indicates that 80.95% of the 231 design interfaces within organizational boundaries are matched by team interactions while 47.93% of the 338 design interfaces across organizational boundaries are matched by team interactions.

$\chi^2_{\text{within-boundaries}} = 0.095$ $\chi^2_{\text{across-boundaries}} = 8.740$ Critical $\chi^2_{(0.99,1)} = 6.635$

4.5.2. Effects due to design interface strength

We define the strength of a design interface by the number and level of criticality of the design dependencies. Hence,

$$\text{design interface strength} = \sum_{i=\text{dependency type}} |c_i| \tag{4.19}$$

where,

dependency type = [spatial, structural, material, energy, information]

c_i = level of criticality of dependency "i" = [-2,-1,0,+1,+2]

Among the 569 design interfaces, 371 of those comprise critical dependencies only (required and/or detrimental dependencies), 39 design interfaces comprise non-critical dependencies only (desired and/or undesired dependencies), and 159 comprise a mixture of both critical and non-critical dependencies.

To perform a chi-square test of independence with respect to hypothesis H3.2, we categorize the 569 design interfaces (YES column of Figure 4.1) according to the following two criteria:

- **First criterion**: Whether a design interface is matched by a team interaction or not.
- **Second criterion**: Whether a design interface is either weak (design interface strength ≤ 4) or strong (design interface strength >4). Note that the average design interface strength is 4.4.

The chi-square test of independence resulted in a $\chi 2$ of 21.385, exceeding the critical value of 6.635 (for one degree of freedom and $\alpha = 0.01$). Hence, we reject the null hypothesis that matching a design interface by a team interaction is independent of the strength of the design interface. More specifically, of the 319 weak design interfaces, 53% were matched by team interactions, whereas of the 250 strong design interfaces 72% were matched by team interactions. Therefore the empirical evidence supports hypothesis H2 that strong design interfaces are more likely to be matched by team interactions than weak design interfaces (see Table A.1).

4.5.2.1. Moderating effects due to systems modularity

We complete chi-square tests of homogeneity to explore whether the effect of design interface strength is statistically different for modular and integrative systems. We found that the effect of design interface strength is homogenous for both modular and integrative systems. The chi-square tests shown in Table A.2 resulted in a χ^2 equal to 0.888 for the cases of weak design interfaces, which is well below the critical value of 6.635 (for α=0.01 and one degree of freedom). Similarly, for the cases of strong design interfaces χ^2 equaled 0.710, which is also below the critical value. Hence, the effect of design interface strength on communication patterns between design teams is not moderated by systems modularity.

4.5.2.2. Moderating effects due to organizational boundaries

Before proceeding further we should test whether the criteria used to categorize design interfaces based on organizational boundaries and design interface strength are independent of each other. We found that the portion of strong design interfaces within organizational boundaries is statistically significant greater than the portion of weak design interfaces within organizational boundaries. Similarly, the portion of weak design interfaces across organizational boundaries is statistically significant greater than the portion of strong design interfaces across organizational boundaries (see Table A.3). The test resulted in a $\chi 2$ of 33.214, exceeding the critical value of 6.635 (for one degree of freedom and $\alpha = 0.01$).

This result suggests we should test the effect of organizational boundaries and design interface strength simultaneously in order to capture the interaction effect between these two hypothesized effects. To do so, we construct a three-dimensional contingency table (see Table 4.3) whose expected values are estimated according to equation 4.8. As stated before, the model specified in 4.8 allows for testing the null hypothesis that the joint effect of design interface strength and organizational boundary is independent of whether such design interface is matched by its corresponding team interaction. The results exhibited in Table 4.3 supports the rejection of the null hypothesis. The test resulted in a $\chi 2$ of 71.463, exceeding the critical value of 11.345 (for three degrees of freedom and $\alpha = 0.01$).

More specifically, under the null hypothesis specified by equation 4.8 the portion of strong design interfaces (for both within and across boundaries) matched by team interactions were statistically significant larger than the portion of weak design interfaces (for both within and across boundaries). ($\chi^2_{\text{within-boundaries}} = 43.175$, $\chi^2_{\text{across-boundaries}} = 28.288$, both greater than the critical value of 6.635 -for one degree of freedom and $\alpha = 0.01$-) Similarly, the portion of design interfaces within organizational boundaries (for both weak and strong interfaces) matched by team interactions were statistically significant larger than the portion of design interfaces across organizational boundaries (for both weak and strong interfaces). ($\chi^2_{\text{weak}} = 30.085$, $\chi^2_{\text{strong}} = 41.378$, both greater than the critical value of 6.635 -for one degree of freedom and $\alpha = 0.01$)

Table 4-3. Joint effects of design interface strength and organizational boundaries †

	Total	Expected cases of design interfaces matched by team interactions	Expected cases of design interfaces not matched by team interactions	Actual cases of design interfaces matched by team interactions	Actual cases of design interfaces not matched by team interactions	$\chi2$ of cases of design interfaces matched by team interactions	$\chi2$ of cases of design interfaces not matched by team interactions
Weak design interfaces (Within organizational boundaries)	96	58.9 (61.34%)	37.1 (38.66%)	69 (71.88%)	27 (28.13%)	1.735	2.749
Weak design interfaces (Across organizational boundaries)	223	136.8 (61.34%)	86.2 (38.66%)	100 (44.84%)	123 (55.16%)	9.894	15.707
Total	319			169	150	**11.629**	**18.456**
Strong design interfaces (Within organizational boundaries)	135	82.8 (61.34%)	52.2 (38.66%)	118 (87.41%)	17 (12.59%)	14.975	23.716
Strong design interfaces (Across organizational boundaries)	115	70.6 (61.34%)	44.4 (38.66%)	62 (53.91%)	53 (46.09%)	1.038	1.649
Total	250			180	70	**16.013**	**25.365**

† Expected values are determined according to equation 4.8, which indicates that 61.34% of the design interfaces on each row of the table are expected to be matched by team interactions.

H_o: The joint effect of design interface strength and organizational boundary is independent of whether the design interface is matched by its corresponding team interaction.

$\chi^2 = 71.463$ Critical $\chi^2_{(0.99,3)} = 11.345$ Since $\chi^2 >$ Critical $\chi^2_{(0.99,3)}$, we reject H_o.

Additionally, we test the null hypothesis that the effect due to organizational boundaries is homogenous throughout the data (for both weak and strong design interfaces). We performed a Chi-square test of homogeneity and found that for the cases within organizational boundaries, the portion of strong design interfaces matched by team interactions was statistically significant greater than the portion of weak design interfaces matched by team interactions ($\chi2 = 8.778$), which is in line with hypothesis H3.2 (see Table 4.4). However, for the cases across organizational boundaries we could not reject the null hypothesis that weak design interfaces are as likely to be matched by team interactions as strong design interfaces ($\chi^2 = 2.501$), which is contrary to hypothesis H3.2.

Table 4-4. Chi-square test of homogeneity. Organizational boundaries controlling for design interface strength.

	Total	Expected cases of design interfaces matched by team interactions	Expected cases of design interfaces not matched by team interactions	Actual cases of design interfaces matched by team interactions	Actual cases of design interfaces not matched by team interactions	$\chi 2$ of cases of design interfaces matched by team interactions	$\chi 2$ of cases of design interfaces not matched by team interactions
Within organizational boundaries (Weak design interfaces)	96	77.714 (80.95%)	18.286 (19.05%)	69 (71.88%)	27 (28.13%)	0.977	4.153
Within organizational boundaries (Strong design interfaces)	135	109.286 (80.95%)	25.714 (19.05%)	118 (87.41%)	17 (12.59%)	0.695	2.953
Total	231	187.000	44.000	187	44	**1.672**	**7.106**
Across organizational boundaries (Weak design interfaces)	223	106.882 (47.93%)	116.118 (52.07%)	100 (44.84%)	123 (55.16%)	0.443	0.408
Across organizational boundaries (Strong design interfaces)	115	55.118 (47.93%)	59.882 (52.07%)	62 (53.91%)	53 (46.09%)	0.859	0.791
Total	338	162.000	176.000	162	176	**1.302**	**1.199**

H_0: The effects of organizational boundaries are the same for weak design interfaces as for strong design interfaces.

$\chi^2_{\text{within-boundaries}} = 8.778$ $\chi^2_{\text{across-boundaries}} = 2.501$ Critical $\chi^2_{(0.99,1)} = 6.635$

We also need to test the null hypothesis that the effect due to design interface strength is homogenous throughout the data (for both within-boundary and across-boundary design interfaces). We found that, for both weak and strong design interfaces, the likelihood that a design interface is matched by a team interaction is greater when it is within organizational boundaries (see Table 4.5). The test resulted in a $\chi 2$ equal to 19.685 for weak design interfaces, and $\chi 2$ equal to 34.558 for strong design interfaces.

Table 4-5. Chi-square test of homogeneity. Design interface strength controlling for organizational boundaries.

	Total	Expected cases of design interfaces matched by team interactions	Expected cases of design interfaces not matched by team interactions	Actual cases of design interfaces matched by team interactions	Actual cases of design interfaces not matched by team interactions	χ2 of cases of design interfaces matched by team interactions	χ2 of cases of design interfaces not matched by team interactions
Weak design interfaces (Within organizational boundaries)	96	50.859 (52.98%)	45.141 (47.02%)	69 (71.88%)	27 (28.13%)	6.471	7.290
Weak design interfaces (Across organizational boundaries)	223	118.141 (52.98%)	104.859 (47.02%)	100 (44.84%)	123 (55.16%)	2.786	3.138
Total	319	169.000	150.000	169	150	**9.256**	**10.429**
Strong design interfaces (Within organizational boundaries)	135	97.200 (72.00%)	37.800 (28.00%)	118 (87.41%)	17 (12.59%)	4.451	11.446
Strong design interfaces (Across organizational boundaries)	115	82.800 (72.00%)	32.200 (28.00%)	62 (53.91%)	53 (46.09%)	5.225	13.436
Total	250	180.000	70.000	180	70	**9.676**	**24.882**

H_o: The effects of design interface strength are the same for within-boundary design interfaces as for across-boundary design interfaces.

$\chi^2_{weak\text{-}interfaces} = 19.685$ $\chi^2_{strong\text{-}interfaces} = 34.558$ Critical $\chi^2_{(0.99,1)} = 6.635$

As a result, we conclude that the effects of organizational boundaries are more severe than the effects of design interface strength. That is, we found empirical support for hypothesis H3.1 throughout the data (for both weak and strong design interfaces). On the other hand, the data support hypothesis H3.2 within organizational boundaries only, while across organizational boundaries design interface strength makes no statistically significant difference on whether or not design interfaces are matched by team interactions.

4.5.3. Effects due to design dependency type

We test the effects due to the type of design dependency by selecting the subset of design interfaces in which only one type of dependency was identified. A total of 122 design interfaces (out of 569 design interfaces) were categorized according to the following two criteria:

- **First criterion**: Whether a design interface is matched by a team interaction or not.
- **Second criterion:** Whether a design interface is either:
  - Spatial-type dependency, or
  - Transfer-type dependency. This category includes structural and/or energy and/or material[4].

Table A.4 shows the results of the chi-square test of independence, which resulted in a $\chi 2$ of 3.225, below the critical value of 3.841 (for one degree of freedom and $\alpha = 0.01$). This result did not allow us to reject the null hypothesis that matching a design interface by team interactions is independent of the type of design interface.

4.5.3.1. Moderating effects due to systems modularity

The subset of 122 design interfaces were split in two subgroups according to the modularity of the systems involved in the design interface. A chi-square test of homogeneity was used to explore hypothesis H3.3. The chi-square test resulted in a $\chi^2_{\text{spatial-type}} = 4.360$ and $\chi^2_{\text{transfer-type}} = 6.035$, which are both greater than the critical value of 3.841 (for α=0.05 and one degree of freedom). These results allow us to reject the null hypothesis that spatial-type design interfaces and transfer-type design interfaces are equally handled by modular design teams and integrative design teams.

The results obtained support the hypothesis that teams designing modular systems have a stronger preference, ability, or willingness, to deal with spatial-type design interfaces than do teams designing integrative systems. As shown in Table A.5, 62.5% of the modular spatial-type design interfaces analyzed were matched by team interactions, while 29.4% of the integrative spatial-type design interfaces were matched by team interactions. Similarly, data also support the hypothesis that teams that design integrative systems are more willing to deal with transfer-type design interfaces than modular design teams. Table A.5 shows that 45.0% of the integrative transfer-type design interfaces analyzed were matched by team interactions, while only 19.5% of the modular transfer-type design interfaces were matched by team interactions.

---

[4] We do not include information-type dependencies in this category because they are only relevant for the externals and controls systems.

4.5.4. Effects due to redesign of the interface

We carried out follow up interviews to better understand qualitatively the reasons for some design interfaces not being matched by team interactions. Some of the reasons that came up during those interviews were that "the interface had not changed with respect to the previous design" and therefore no interaction needed to take place, which is the argument hypothesized in H3.4a.

To test the effect of redesign of the interface many factors need to be taken into account. We need to assess not only how much the engine model changed with respect to the previous generation but also how much the development organization changed with respect to the organization involved in the design of previous generations (as hypothesized in H3.4b). Therefore, hypothesis H3.4 is directly related to the dynamic problem of how the product architecture and the development organization evolve with time, which is in our agenda for future research.

We completed limited statistical tests based on the assumption that the development organization which developed the engine studied had access to the knowledge learned in the design of previous generations. The results of this analysis, presented in Appendix B, do not strongly support hypothesis H3.4. That is, there is not statistically significant difference in the way the unchanged design interfaces were addressed by team interactions with respect to the design interfaces that exhibited certain percentage of redesign.

4.5.5. Effects due to potential indirect team interactions through other design teams

Based on a fundamental property of the adjacency matrix from graph theory (Harary (1969), Gebala and Eppinger (1991)) we square the 54x54 binary team interaction matrix to obtain the number of potential indirect interactions between each pair of teams. Since we defined potential indirect team interactions for the cases when no direct interactions were reported, we analyzed the 2439 cases corresponding to the NO row of Figure 4.1. Hence,

$$
\begin{aligned}
&K^{o}_{54,54} = \textit{Potential Indirect Interaction Matrix (through other design teams)}\\
&K^{o}_{54,54} = \mathrm{T}^{b}_{54,54} \times \mathrm{T}^{b}_{54,54} \quad \textit{where } \mathrm{T}^{b}_{54,54} = \textit{Team Interaction Matrix (binary) from (2.2)}\\
&k^{o}_{i,j} = \textit{Number of potential indirect team interactions from team "i" to team "j"}\\
&k^{o}_{i,j} \textit{ is undefined for } t^{b}_{i,j} = 1, \textit{ and for } i = j
\end{aligned}
\qquad (4.20)
$$

The sample analyzed was categorized according to the following two criteria:

- **First criterion**: Whether or not a pair of teams shares a design interface (NO row of Figure 4.1).
- **Second criterion**: The number of potential indirect team interactions through other design teams:
  - No potential indirect team interactions
  - One potential indirect team interaction
  - More than one potential indirect team interaction

The chi-square test of independence resulted in a $\chi 2$ of 93.911, far exceeding the critical value of 9.210 (for two degrees of freedom and $\alpha = 0.01$) (see Table A.6). More specifically, 5.0% of the cases with no potential indirect team interactions coincided with design interfaces, 10.2% of the cases with one potential indirect team interaction coincided with design interfaces, and 19.8% of the cases with more than one potential indirect team interaction coincided with design interfaces. These results support hypothesis H3.5 that teams, which share a design interface but do not interact directly are more likely to interact indirectly through other design teams.

4.5.5.1. Moderating effects due to systems modularity

We complete a chi-square test of homogeneity to explore whether the effect due to potential indirect team interactions thorough other design teams is statistically different for modular and integrative systems. We found that such effect is not statistically significant for both modular and integrative systems. The chi-square tests shown in Table A.7 resulted in an overall χ^2 equal to 11.939, which is below the critical value of 15.086 (for α=0.01 and five degrees of freedom).

4.5.5.2. Moderating effects due to organizational boundaries

We next test whether the effects of indirect team interactions (through other design teams) are independent of organizational boundaries. We found that a statistically significant higher number of potential indirect team interactions occur within organizational boundaries (test resulted in a $\chi 2$ of 94.529, exceeding the critical value of 9.210 (for two degrees of freedom and $\alpha = 0.01$) (see Table A.8). Therefore, we need to test if the effects due to indirect team interactions still hold when controlling for organizational boundaries.

After splitting the sample in two subgroups, within and across organizational boundaries, we complete separated tests of independence. The results of the chi-square tests supported hypothesis H3.5 for the cases across organizational boundaries (test shown in Table A.9 resulted in a $\chi 2$ of 81.876, exceeding the critical value of 9.210 for two degrees of freedom and $\alpha =$ 0.01). More specifically, 17.3% of the cases with more than one potential indirect interaction coincided with design interfaces, whereas 9.5% of the cases with one potential indirect interactions coincided with design interfaces, and only 4.0% of the cases with no potential indirect interactions coincided with design interfaces. For the cases within organizational boundaries, the data did not support hypothesis H3.5 (the test resulted in a $\chi 2$ of 6.104) (see Table A.9).

The results in Table A.9 are confirmed by the chi-square test based on the model specified in equation 4.8 (see Table A.10). Table A.10 shows the results of testing the null hypothesis that the joint effect of potential indirect interactions through other design teams and organizational boundaries is independent of whether such case would coincide with a design interface. Even though the test resulted in a $\chi 2$ of 129.038 for the cases within boundaries, there is not statistically significant association between the number of potential indirect team interactions and the presence of design interfaces. However, in the cases of across organizational boundaries the test not only resulted in a $\chi 2$ of 75.757 (exceeding the critical value of 9.210 for two degrees of freedom and $\alpha = 0.01$), but also exhibited a statistically significant association between the number of potential indirect team interactions and the presence of a design interface. This supports the hypothesis H3.5.

4.5.6. Indirect team interactions through system integration teams

In order to determine the number of potential indirect team interactions through system integration teams, we first squared the 60x60 team interaction binary matrix to obtain the potential indirect teams interactions both through other design teams and through system integration teams. We took the first 54x54 portion of this matrix and subtracted the 54x54 matrix that contained the potential indirect teams interactions through other design teams, resulting in a matrix that contains the number of potential indirect interactions through system integration teams only.

$K^{s}_{54,54} = \textit{Potential Indirect Team Interaction Matrix (through system integration teams only)}$

$K^{s}_{54,54} = K^{t}_{54,54} - K^{o}_{54,54}$

where, $K^{o}_{54,54} = \textit{Potential Indirect Team Interaction Matrix (through other design teams only) from (4.20)}$

$$K^{t}_{60,60} = \textit{Potential Indirect Team Interaction Matrix} \tag{4.21}$$

(through both other design teams and system integration teams)

$K^{t}_{60,60} = T^{b}_{60,60} \times T^{b}_{60,60}$ *where,* $T^{b}_{60,60} = \textit{Team Interaction Matrix (binary) from (2.2)}$

$K^{t}_{54,54} = I_{54,60} \times K^{t}_{60,60} \times I_{60,54}$ where, $I_{54,60}$ *and* $I_{60,54}$ *contain "1s" in their diagonal and "0s" otherwise.*

We analyzed a sample of 2439 cases in which no direct team interactions were reported (NO row of Figure 4.1). Similar to the case of potential indirect interactions through other design teams, we categorized the data according to the following two criteria:

- **First criterion**: Whether or not a pair of teams shares a design interface.
- **Second criterion**: The number of potential indirect team interactions through system integrators:
  - No potential indirect team interactions
  - One potential indirect team interaction
  - More than one potential indirect team interaction

The Chi-square test of independence performed resulted in a $\chi 2$ of 1.435, well below the critical value of 9.210 for two degrees of freedom and $\alpha = 0.01$ (see Table A8). Hence, we accept the null hypothesis that the number of potential indirect interactions through system integration teams is independent of whether or not those teams share a design interface (contrary to hypothesis H3.5).

4.6. Team interactions Not Predicted by Design Interfaces

In this section, we test the hypotheses that help us to understand the cases when team interactions were reported even though such teams did not share known design interfaces (the 74 "O" cells of Figure 4.1).

4.6.1. Effects due to system boundaries

We hypothesize in H3.6 that unknown design interfaces are more likely to occur across system boundaries than within system boundaries, resulting in a higher portion of unpredicted team interactions across system boundaries than within system boundaries. To test this hypothesis, we select the 423 team interactions (YES row of Figure 4.1), and categorize them according to the following two criteria:

- **First criterion**: Whether or not a team interaction is predicted by a design interface.
- **Second criterion**: Whether the design interface corresponding to each pair of teams is within or across system boundaries.

The chi-square test of independence performed resulted in a $\chi 2$ of 15.517, exceeding the critical value of 6.635 (for one degree of freedom and $\alpha = 0.01$). More specifically, of the 208 team interactions within system boundaries, 89.9% were predicted by design interfaces whereas 75.4% of the 215 team interactions across system boundaries were predicted by design interfaces These results, shown in Table A.12, support hypothesis H3.6.

4.6.1.1. Moderating effects due to systems modularity

We want to test whether there is a statistically significant difference in the way design interfaces predict team interactions of modular versus integrative design teams. The chi-square tests of homogeneity shown in Table A.13 resulted in a χ^2 equal to 0.482 for the cases within system boundaries, which is well below the critical value of 6.635 (for α=0.01 and one degree of freedom). On the other hand, for the team interactions across system boundaries the χ^2 equaled 9.567, which is above the critical value.

These results allow us to accept the null hypothesis that the portion of unknown design interfaces within system boundaries is equivalent for both modular and integrative systems. However, for interactions across system boundaries we found that the portion of unknown design interfaces across modular systems is statistically significant larger than for integrative systems. Note that unknown design interfaces are characterized by unpredicted team interactions ("O" cells). Specifically, 38.5% of the modular design team interactions across system boundaries were not predicted by design interfaces whereas only 18.7% of the integrative design team interactions across system boundaries were not predicted by design interfaces.

4.6.2. Effects due to potential secondary design interfaces

In order to quantify the effect of system-level design dependencies, we square each of the binary matrices corresponding to each design dependency type to obtain the number of secondary design dependencies of each dependency type. The total number of potential secondary design dependencies is obtained by adding up the five squared matrices. We now have a mapping of the potential secondary design interfaces for the 2293 cases in which no direct design interface was known (NO column of Figure 4.1). In algebraic terms we have the following:

$$
\begin{aligned}
&S^{t}_{54,54} = \textit{Secondary Design Interface Matrix}\\
&S^{t}_{54,54} = \sum_{k=\textit{dependency type}} S^{k}_{54,54}\\
&s^{t}_{i,j} = \textit{Total number of secondary dependencies on which component } j \textit{ depends on component } i\\
&\text{where,}\\
&S^{k}_{54,54} = A^{k}_{i,j} \times A^{k}_{i,j}\\
&\text{where,}\\
&a^{k}_{i,j} = 1 \textit{ if component "} j \textit{" depends on component "} i \textit{" (dependency "} k \textit{" type)}\\
&a^{k}_{i,j} = 0 \textit{ otherwise}
\end{aligned}
\tag{4.22}
$$

The sample analyzed is formed by the 2293 cases where potential secondary design dependencies are defined (NO column of Figure 4.1), and it is categorized according to the following two criteria:

- **First criterion**: Whether or not team interaction was reported.
- **Second criterion**: The number of secondary design dependencies (Note that the average number of secondary design dependencies is 2.4):
  - No secondary design dependency
  - Three or fewer secondary dependencies
  - More than three secondary dependencies

We found empirical evidence supporting hypothesis H3.7. The test resulted in a $\chi 2$ of 51.561, far exceeding the critical value of 9.210 (for two degrees of freedom and $\alpha = 0.01$). Specifically, 0.5% of the cases with no secondary design dependencies were matched by team interactions, whereas 3.9% of the cases with three or fewer secondary dependencies were matched by team interactions, and 6.8% of the cases with more than three secondary dependencies were matched by team interactions (see Table A.14).

4.6.2.1. Moderating effects due to systems modularity

We complete a chi-square test of homogeneity to explore whether the effect due to potential secondary dependencies is statistically difference for modular and integrative systems. We found that such effect is not statistically significant for both modular and integrative systems. The chi-square tests shown in Table A.15 resulted in an overall χ^2 equal to 9.717, which is below the critical value of 15.086 (for α=0.01 and five degrees of freedom).

4.6.2.2. Moderating effects due to system boundaries

We found that a higher number of secondary design dependencies occur within system boundaries (test, shown in Table .16, resulted in a $\chi 2$ of 97.796, exceeding the critical value of 9.210 for two degrees of freedom and $\alpha = 0.01$). Therefore, we need to test if the effects due to secondary design dependencies still hold when controlling for system boundaries.

After splitting the sample in two subgroups, within and across system boundaries, we complete separated tests of independence. The results of the chi-square tests shown in Table A.17 support hypothesis H3.7 for the cases across system boundaries (test resulted in a $\chi 2$ of 37.203, exceeding the critical value of 9.210 for two degrees of freedom and $\alpha = 0.01$). However, the data did not support hypothesis H3.7 for the cases within system boundaries (test resulted in a $\chi 2$ of 1.304).

The results in Table A.17 are confirmed by the chi-square test based on the model specified in equation 4.8 (see Table A.18). Table A.18 shows the results of testing the null hypothesis that the joint effect of potential secondary design dependencies and system boundaries is independent of whether such case would coincide with a team interaction. Even though the test resulted in a $\chi 2$ of 112.489 for the cases within boundaries, there is not statistically significant association between the number of potential secondary design dependencies and the presence of team interactions. In contrast, for the cases across organizational boundaries the test not only resulted in a $\chi 2$ of 32.713 (exceeding the critical value of 9.210 for two degrees of freedom and $\alpha = 0.01$), but also exhibited a statistically significant association between the number of potential secondary design dependencies and the presence of team interactions, supporting hypothesis H3.7.

A note of caution regarding the results for the cases within boundaries exhibited in Tables A.17 and A.18 needs to be made. The small expected frequencies in some of the cases within boundaries poses a possible threat to validity of the chi-square tests shown in these tables. Statisticians disagree as to how to handle this problem. We will follow the recommendation of Cochran (1952, 1954), who states that for tables with more than one degree of freedom a minimum expected frequency per cell of 1 is permissible if no more than 20% of the cells have expected frequencies of less than 5. In any case, the tests provide some empirical evidence to reject hypothesis H3.7 for within-boundary cases.

4.7. Discussion of Results

Table 4.6 provides a summary of the results obtained from testing the hypotheses posed in Section 3. It is important to note that there are other ways to explain the unexpected cases in our data set. One important explanation we could not test is the presence of measurement errors. We recognize that not all known design interfaces were captured during our data collection, nor were all coordination-type interactions reported during the survey. We also recognize that some design interfaces might have been mistakenly documented and some team interactions might have been erroneously reported.

We first focused our analysis on the effects imposed by organizational boundaries, design interface strength, design interface redesign, and indirect team interactions to explain the cases when design interfaces were not matched by team interactions. We found that design interfaces across organizational boundaries are less likely to be matched by team interactions (in line with hypothesis H3.1). We also found that weak design interfaces are less likely to be matched by team interactions than strong design interfaces (in line with hypothesis H3.2).

We analyzed the cases with no direct team interaction to study the effects of potential indirect team interactions through other design teams. We found that the portion of cases that do not interact directly, but share a design interface, was statistically significant higher in the presence of higher number of potential indirect team interactions, which is in line with hypothesis H3.5. We found no empirical support of this effect for the teams within organizational boundaries (contrary to hypothesis H3.5). This suggests that teams within a group interact directly to handle their design interfaces, whereas teams across groups might not interact directly because they obtain the information needed through other design teams (perhaps, even from their same group).

We also studied the effects of potential indirect team interactions through system integration teams (the last six teams of the team interaction matrix shown in Figure 2.6). We did not find statistically significant evidence that supports the hypothesis that a higher number of potential indirect interactions through system integration teams is associated with teams that share a design interface (hypothesis H3.5). This result corresponds with the nature of the work performed by system integration teams. These teams interact with design teams in every group of the development organization making the likelihood of potential indirect team interaction (through these teams) independent of whether or not two teams share a design interface.

We also studied the effects of system boundaries, and the extent to which secondary design dependencies explain the cases when team interactions were not predicted by design interfaces. We found that team interactions across system boundaries were less likely to be predicted by design interfaces (in line with hypothesis H3.6). That is, design experts are more likely to fail to identify design interfaces across system boundaries than within system boundaries.

We studied the effects of secondary design dependencies by analyzing the cases where no direct design interfaces were identified. We found, for the cases across system boundaries, that the higher the number of secondary dependencies the more likely the teams involved in those design interfaces interact (in line with hypothesis H3.7). However, the data did not support hypothesis H3.7 for the cases within system boundaries. This result implies that higher-order dependencies are not significant within system boundaries. Indeed, the system-level dependencies are expected to occur between components of different systems, and some are handled by the design teams themselves.

Table 4-6. Summary of results

<table>
<tr>
<td colspan="2">8% of the cases were design interfaces not matched by reported team interactions.
Explanation for these cases to occur (statistically supported) include:
• Design interfaces across organizational boundaries (supporting hypothesis H3.1).
• Weak design (supporting hypothesis H3.2).
• Unexpected design interface type (supporting hypothesis H3.3)
• Indirect interactions through other design teams across organizational boundaries (partially supporting hypothesis H3.5).</td>
<td>78% of the cases revealed no design interfaces and no corresponding team interactions.
These cases were expected to occur.</td>
</tr>
<tr>
<td>12% of the cases revealed design interfaces matched by team interactions.
These cases were expected to occur.</td>
<td colspan="2">2% of the cases were team interactions not predicted by design interfaces.
Explanations for these cases to occur (statistically supported) include:
• Unknown design interfaces occur across system boundaries (supporting it hypothesis H3.6).
• Higher number of secondary design interfaces across system boundaries (partially supporting hypothesis H3.7).</td>
</tr>
</table>

4.7.1.1. Designing Modular versus Integrative Systems

We analyze the effects of system modularity whose results can be summarized as follows:

- The statistically significant differences in the way integrative design teams handle design interfaces across boundaries suggest that these teams are more efficient at overcoming the barriers imposed by organizational boundaries. The distributed nature of the integrative systems forces these design teams to overcome organizational barriers in order to handle design interfaces with all the systems.
- When analyzing the effects of system boundaries, we found a statistically significant larger proportion of unpredicted team interactions ("O" cells of Figure 9) associated with modular systems. Since unpredicted team interactions represent unrecognized design interfaces, we conclude that design interfaces across modular systems are more difficult for design experts to recognize than interfaces with integrative systems.

- The existence of various types of design interfaces and the statistically significant difference in the way they were handled by modular and integrative design teams provide empirical support to the notion of "communication filters" introduced by Henderson and Clark (1990). We found that spatial-type design interfaces are largely addressed in the design of modular systems while transfer-type design interfaces are more likely to be handled in the design of integrative systems.

4.8. Limitations of the Analysis

The analysis presented in this chapter is based on classical techniques used in categorical data analysis. We have based our analysis on the assumption that both the design interface matrix and the team interaction matrix follow a Bernoulli probability distribution (statistically independent cells). However, research in social science has shown that social network data (such as that presented in the team interaction matrix) possess strong deviation from randomness. More specifically, empirical evidence (Holland and Leinhardt, 1981) shows that social networks exhibit several types of dependence: tendency toward *reciprocation*, tendency toward *expansiveness* (i.e. to generate interactions) and tendency toward *attraction* (i.e. to attract interactions).

In the next chapter we present a log-linear model built upon social network analysis research that allows us to validate the results of some of the tests presented in this chapter.

5. *A Log-linear Model*

"Los beneficios que se hacen hoy, se reciben mañana, porque Dios premia la virtud en este mundo mismo"
Simón Bolívar

We address some of the limitations of the categorical data analysis presented in the previous chapter, by developing a log-linear model based on research completed in the area of social network analysis. The models described in this chapter contain parameters that quantify the "structural effects" present in a network, such as reciprocity and tendencies toward differential expansiveness and differential attraction. The objective of this chapter is to illustrate how to validate the assumptions made in the analysis presented in chapter four. The model presented here allows us to test the effects of organizational/system boundaries, and the moderating effects of system modularity. These models are *dyadic interaction* models, which use the natural log of probabilities as their basic modeling unit. Specifically, we will estimate a model of the form:

> *ln* [P(component *i* depends on component *j* and team *i* reports interaction with team *j*)] = F(overall mean, tendency of component *i* to generate design interfaces to other components, tendency of component *j* to depend upon other components, overall tendency to reciprocate design interfaces, tendency of team *i* to report interaction with other teams, tendency of other teams to report interaction with team *j*, overall tendency to reciprocate team interactions, overall association between design interfaces and team interactions, effect due to system/organizational boundaries, effect due to systems modularity) (5.1)

Statistical network analyses allow us to assess a model by measuring the fit of the model to data. Hence, we will compare the observed effects to hypothesized effects, as well as significance tests to determine whether an effect is due to sampling variation (Wasserman and Faust, 1994).

Before we describe our log-linear model, we review some basic statistical and graph-theory notation. We then introduce the Holland and Leinhardt's (1981) p_l distribution. Then, we show how these types of models can be fitted to data by using maximum-likelihood procedures employed to fit log-linear models to multidimensional contingency tables (Fienberg and Wasserman (1981)). Following the work by Fienberg, Meyer and Wasserman (1985) on solving the problem of analyzing multiple relationships in a network, and Wasserman and Iacobucci (1988) on analysis of sequential network data, we propose a base model to describe the network associated with the resultant matrix (Figure 3.X). Finally, following a similar approach to the one presented by Van den Bulte and Moenaert (1998), we extend the base model with structural parameters that capture the effects associated to organizational/system boundaries and with the strength of design interface, so that some of the hypotheses could be formally tested.

5.1. Mathematical Notation

Let D_g be a specific digraph on g nodes from a single binary relation R with at most one arc connecting node i to node j. In our context, the mathematical term *node* refers to a component of the digraph associated with the design interface matrix or a team of the digraph associated with the team interaction matrix. The mathematical term *arc* refers to the presence of a design interface between components, or a team interaction between two design teams. Each digraph has a corresponding adjacency matrix $\mathbf{X}$ with elements (X_{ij}). The binary design interface matrix and the binary team interaction matrix are the adjacency matrices of their corresponding digraphs.

Since our unit of analysis is the dyad, we need to provide some mathematical notation for it. We will label a dyad as follows:

$D_{ij} = (X_{ij}, X_{ji})$ for $i<j$

Hence, D_{ij} is a bivariate random variable with $2^2 = 4$ possible states. These states are:

$D_{ij} = (1,1)$; if *mutual* interaction

$D_{ij} = (1,0)$ or $(0,1)$; if *asymmetric* interaction

$D_{ij} = (0,0)$; if *null* interaction

Also, let us define the following quantities associated to adjacency matrices:

$M = \sum_{i<j} X_{ij} \cdot X_{ji}$; hence, M = number of mutual interactions

$X_{i+} = \sum_{j=1}^{g} X_{ij}$; *out-degree* of node i, which is the total number of non-zero entries in row i

$X_{+j} = \sum_{i=1}^{g} X_{ij}$; *in-degree* of node j, which is the number of non-zero entries in column j

A detailed discussion of these and other related network summaries can be found in Harary, *et al* (1965).

Fienberg and Wasserman (1981) introduced an alternative notation to describe the basic structure for the study of dyads with a single relation. The objective of this alternative notation is to facilitate the analysis of the network as a categorical data set for which standard statistical software packages provide solutions. Let us consider a four-dimensional $g \times g \times 2 \times 2$ cross-classification (the **Y**-array), $\mathbf{Y} = [Y_{ijkl}]$, where the subscripts i and j refer to the two elements in a dyad while k and l refer to the state of the dyad, hence

$Y_{ijkl} = 1$ if $D_{ij} = (k,l)$

$Y_{ijkl} = 0$, otherwise

Therefore, for a given dyad (i,j) we obtain a 2 x 2 table as shown in Table 5.1. It is important to note that $Y_{ij00} + Y_{ij10} + Y_{ij01} + Y_{ij11} = 1$ for all $(i \neq j)$, hence these 2 x 2 tables contain one 1 and three 0's. Also, $Y_{ijkl} = Y_{jilk}$, and thus we should only consider $i<j$, however we will retain the entire **Y**-array to facilitate the process of fitting the models to data.

$l = X_{ji}$	1	Y_{ij01}	Y_{ij11}
	0	Y_{ij00}	Y_{ij10}
		0	1
		$k = X_{ij}$	

Table 5-1. Dyadic Contingency Table

5.2. The p_1 Distribution

There are empirical results from social science that support the notion that social adjacency matrices, such as our team interaction matrix, are not distributed randomly and that, in fact, they exhibit expected non-random behavior. Can we say the same for adjacency matrices that contain relations between the components of a product, such as our design interface matrix? We believe so, and we based the analysis presented in this chapter on this fundamental hypothesis. Indeed, we test this hypothesis when evaluating the goodness-of-fit of our models.

Holland and Leinhardt (1981) developed a model termed p_l to emphasize that is the first or simplest family distribution that expresses the three elementary social tendencies of *reciprocation, differential expansiveness* and *differential attraction.*

The model discussed in this section assigns probability functions on the links between elements (i.e. components or design teams) by specifying the probability that a pair of elements has one of four possible dyadic relationships (see Table 5.1). The entire network of g ($g = 54$) elements is decomposed into an equivalent set of $\binom{g}{2}$ dyads. To determine the probability distribution of the network, the dyads are assumed to be conditionally independent, so that we just multiply the dyad probability distributions to obtain their joint probability distribution.

Regarding the independent dyads assumption Holland and Leinhardt (1981) pointed out that "this independence assumption means that p_l cannot express tendencies toward transitivity, cliquing, hierarchy, and so on, other than those already implied by tendencies toward reciprocation and differential attraction". On the same issue, Fienberg, *et al* (1985) stated that "the assumption of independent dyads is common to many recent models for networks, although it is, at best, an approximation to reality. But building into the models either a dependence structure among dyads or probability distributions on larger subgraphs such as triads appears to be very difficult".

Holland and Leinhardt (1978) present empirical evidence that the independence dyads assumption may be satisfied in a significant number of social network studies. Hence, the p_l distribution may not only provide a null model but also provide adequate models for representing certain types of empirical data associated to networks.

Given the **Y**-array, the Holland and Leinhardt distribution, p_l, can be expressed as follows:

$$\ln \mathrm{P}\{Y_{ij00} = 1\} = \lambda_{ij}$$
$$\ln \mathrm{P}\{Y_{ij10} = 1\} = \lambda_{ij} + \theta + \alpha_i + \beta_j$$
$$\ln \mathrm{P}\{Y_{ij01} = 1\} = \lambda_{ij} + \theta + \alpha_j + \beta_i$$
$$\ln \mathrm{P}\{Y_{ij11} = 1\} = \lambda_{ij} + 2\theta + \alpha_i + \beta_i + \alpha_j + \beta_j + \rho$$

or in shorthand,

$$\ln \mathrm{P}\{Y_{ijkl} = 1\} = \lambda_{\mathrm{ij}} + (k+l)\theta + k \cdot \alpha_{\mathrm{i}} + l \cdot \beta_{\mathrm{i}} + l \cdot \alpha_{\mathrm{j}} + k \cdot \beta_{\mathrm{j}} + (kl)\rho \qquad (5.2)$$

subject to the constraints

$$\mathrm{P}\{Y_{ij00} = 1\} + \mathrm{P}\{Y_{ij10} = 1\} + \mathrm{P}\{Y_{ij01} = 1\} + \mathrm{P}\{Y_{ij11} = 1\} = 1 \qquad (5.3)$$

for all dyads, and

$$\sum_{\mathrm{i}=1}^{g} \alpha_{\mathrm{i}} = \sum_{\mathrm{j}=1}^{g} \beta_{\mathrm{j}} = 0 \qquad (5.4)$$

The parameters $\{\alpha_i\}$ measure the *expansiveness* or "productivity" of the elements of the network, indicating how likely an element is to generate relational ties (non-zero cells in row *i* of the matrices). The parameters $\{\beta_j\}$ measure the *attraction* or "popularity" of the elements of the network, indicating how likely an element is to receive relational ties (non-zero cells in column *j* of the matrices). The "reciprocity" parameter, ρ, measures the overall tendency in the network to reciprocate interactions. The θ parameter indicates the overall volume of interaction in the network. Finally, the λ_{ij} parameters are "dyadic" effects that ensure that the probabilities sum to one for each dyad (equation 5.3), they have no substantive meaning. For a more detailed description of these parameters refer to Holland and Leinhardt (1981).

5.3. Fitting p_1 to a Single Relational Data

Fienberg and Wasserman (1981) presented the log likelihood function for the model p_1 given the **Y**-array as follows:

$$log\ L(\{\alpha_i\},\{\beta_j\},\{\lambda_{ij}\},\rho,\theta \mid \mathbf{Y}) = (\rho/2)\cdot y_{++11} + \sum_{i=1}^{g} \alpha_i \cdot y_{i+1+} + \sum_{j=1}^{g} \beta_j \cdot y_{+j1+} + \theta \cdot y_{++1+} + \sum_{i<j} \lambda_{ij} \quad (5.5)$$

Note that even though the log likelihood is defined for *i<j*, it can be described through the use of marginal totals of the complete **Y**-array.

Fitting the p_1 model to an observed matrix, such as the binary design interface matrix or the binary team interaction matrix, is equivalent to constructing an "expected" matrix with in-degrees, out-degrees, number of mutuals (M), and total number of choices identical to those of the observed matrix. We then ask how much the expected and observed matrices differ. A large difference indicates that we need to incorporate additional parameters to capture structural effects into the model.

Fienberg and Wasserman (1981) showed that by using the redundant representation of the full **Y**-array, one can transform the statistical problem of fitting p_1 to **X** into an equivalent statistical problem, fitting the "no-three-factor interaction" log linear model to the **Y**-array. Therefore, we can use standard iterative proportional fitting computer programs for contingency tables (we used SPSS), and no need to do any programming to fit p_1.

Following the well-known Fienberg notation (see Fienberg 1980 for details), the no-three-factor interaction log-linear model simultaneously fits the following margins of **Y**:

[12][13][24][14][23][34] (5.5)

The numbers in brackets in (5.5) are the margins of **Y**, which are sufficient statistics for the parameters in the basic model. This notation uses marginal totals as used to describe log linear models that fit to multiway arrays. For example, [12] refers to the two-way marginal totals of the **Y**-array corresponding to the first two subscripts, i and j, that is, [12] corresponds to the margin $\{Y_{ij++}\}$. Because the log-linear models considered here are hierarchical, the 1-dimensional margins for variables 1 and 2, $\{Y_{i+++}\}$and $\{Y_{+j++}\}$, are also included. Similarly, [24] refers to the two-way marginal totals corresponding to the second and fourth subscripts, j and l. The inclusion of the [12] margin assures us that the probabilities add to one for each dyad, that is, it includes the $\{\lambda_{ij}\}$ parameters in the model. The [13] and [24] margins are identical and allow for the inclusion of $\{\alpha_i\}$ and θ in the model. Similarly, the [14] and [23] identical margins correspond to the inclusion of $\{\beta_j\}$. Finally, the [34] margin allows for the inclusion of ρ (for more details refer to Wasserman and Faust, 1994, pp. 625).

After fitting the p_1 model, we need to evaluate how good the model fits the observed data. We compute likelihood-ratio goodness-of-fit statistics according to the following expression provided by Fienberg and Wasserman (1981):

$$G^2 = 2\sum_{i<j}\sum_{k,l} y_{ijkl} \cdot \ln(y_{ijkl} / \hat{y}_{ijkl}) \qquad (5.6)$$

where $\hat{y}_{ijkl}$ = predicted value of Y_{ijkl}.

When calculating the G^2 associated to p_1 from commercial statistical packages we divide the computed G^2 by 2 to adjust for the duplication of $Y_{ijkl}=Y_{jilk}$. To complete an informal goodness-of-fit test using G^2, we need to determine the number of degrees of freedom associated to p_1 to informally compare G^2 with the proper χ^2 distribution. Our test is informal because the sparseness of the **Y**-array does not allow us to use standard asymptotic theory. More specifically, the natural asymptotic setting for this type of network problem is one in which the size of the network increases (without upper bound) with the number of nodes, g, which prevents us to use standard asymptotic theory. Our final model addresses the limitation due to lack of asymptotic theory by aggregating the elements of the network according to external attributes.

Even though cannot formally test goodness-of-fit of a p_1 model, we can complete conditional test statistics, based on Haberman's (1977) results, to check whether all the parameters in the model are required. (Refer to Fienberg and Wasserman (1981) for more details.)

We have to be careful when estimating the number of degrees of freedom (DF) associated with p_1. The general rule is

DF = Number of cells in the contingency table – Number of parameters fitted

There are 2 DF for each 2 x 2 table contained in the **Y**-array; hence $2\binom{g}{2} = g(g\text{-}1) = 2862$ DF for modeling the **Y**-array, which is exactly the number of off-diagonal entries in **X**. When fitting p_1, we use 1 DF for each θ and ρ and $(g\text{-}1)$ DF each for $\{\alpha_i\}$ and $\{\beta_j\}$. Hence, we have a total of $g(g\text{-}3) = 2754$ DF associated to the p_1 model of each of the matrices, the design interface matrix and the team interaction matrix.

The intention of this section was not to build a typical p_1 model for our matrices, but to introduce the fundamental concepts supporting the formulation of the p_1 model, which is the basis of our model.

5.4. Our Base *log-linear* Model

When developing our base log-linear model we follow the following steps:

- Extend the p_1 model to a network with two relations, that is, a network of elements (i.e. design teams) whose designing components may share design interfaces, and the teams themselves may report technical interactions with each other.
- Aggregate components and teams into groups, according to the system boundaries of the product and the organizational boundaries of the development organization.
- Extend the model with association parameters that capture the basic correlation between design interfaces and team interactions.
- Estimate, and compute goodness-of-fit of the base model.

In the next section we describe how we extend the base model to incorporate structural parameters to test the hypotheses related to the effects of organizational and system boundaries, and the moderating effects due to systems modularity.

5.4.1. Extending p_1 to a two-relation data

Fienberg, *et al* (1985) first addressed the problem of extending p_1 to multiple sociometric relations. Wasserman and Iacobucci (1988) used their results as the basis to study sequential network data. More recently, Van den Bulte and Moenaert (1998) used these results to analyze the interactions between R&D teams in two points in time (before and after collocation). We want to extend this approach to develop a base model of the resultant matrix containing both the design interfaces and the team interactions between the 54 elements of our network data.

We expand the p_1 model described in the previous section by considering the joint distribution of both design interfaces and team interactions for a given dyad. That is, each dyad (i,j) consisting of elements i and j will have four (2 x 2) states associated to their design interface relation, times four (2 x 2) states associated to their team interaction relation, resulting in 16 states for each dyad.

We assign the subscripts (k_1,l_1) to describe the states associated to the design interface relation, while the subscripts (k_2,l_2) refer to the states associated to the team interaction relation of dyad (i,j). The redefined **Y**-array has six dimensions 54 x 54 x (2 x 2) x (2 x 2), and its characteristic element can be defined as follows:

$Y_{ij\,k1,l1\,k2,l2} = 1$ if dyad (i,j) behaves as described by (k_1,l_1) for their design interfaces and by (k_2,l_2) for their team interactions.

$Y_{ij\,k1,l1\,k2,l2} = 0$ otherwise.

Considering the joint distribution of design interfaces and team interactions yields to the following log-linear model, which describe simultaneously the behavior of the elements of our network according to two independent relations (design interfaces and team interactions). The log-linear model can be written as follows:

$$\ln P\{Y_{ij\,k_1l_1\,k_2l_2} = 1\} = \lambda_{ij} + (k_1 + l_1)\theta_1 + k_1\alpha_{1i} + l_1\beta_{1i} + l_1\alpha_{1j} + k_1\beta_{1j} + (k_1l_1)\rho_1 + (k_2 + l_2)\theta_2 + k_2\alpha_{2i} + l_2\beta_{2i} + l_2\alpha_{2j} + k_2\beta_{2j} + (k_2l_2)\rho_2 \quad (5.7)$$

The parameters on this model have the same meaning as in the original p_1 model, but applied to either design interfaces (subscript 1) or team interactions (subscript 2).

5.4.2. Aggregating components and teams into groups (The W-array)

Since the elements of our network are categorized into subsets according to their membership to engine's systems or teams' groups, we can aggregate them without using relational information (i.e. design interfaces or team interactions). In other words, the way we aggregate components and teams into groups is not the result of manipulating the data such as clustering or partitioning, it is just the result of inherent attributes of the elements of the network studied.

Fienberg and Wasserman (1981) introduced the approach of placing actors into subsets using relevant actor characteristics so that, actors within a subset are assumed to behave similarly. This assumption of comparable behavior of elements within subsets has been termed *stochastic equivalence* (Wasserman and Weaver, 1985). Assuming that elements i and j are stochastic equivalent means, in mathematical terms, that:

$$\alpha_i = \alpha_j \text{ and } \quad \beta_i = \beta_j \quad (5.8)$$

We operationalize the concept of stochastic equivalence by aggregating the 54 elements of the **Y**-array into 8 subsets[5]. By doing so, we obtain a much smaller **W**-array whose dimensions are 8 x 8 x (2 x 2) x (2 x 2), with elements $\{W_{rs\,k1,l1\,k2,l2}\}$ defined as follows:

[5] We lump the last 10 elements corresponding to External and Controls into one subset.

$$w_{rs\,k_1l_1\,k_2l_2} = \sum_{i \in G_r} \sum_{j \in G_s} y_{ij\,k_1l_1\,k_2l_2} \qquad (5.9)$$

That is, we simply sum the dyads between groups r (G_r) and s (G_s) that behave as (k_1,l_1,k_2,l_2). Note that the entries of the **W**-array are not binaries. Indeed, the minimum value $W_{rs\,k1,l1\,k2,l2}$ can take is zero and the maximum possible value is the number of dyads between groups r and s. The expected value of $W_{rs\,k1,l1\,k2,l2}$ can be expressed as:

$$E(W_{rs\,k1,l1\,k2,l2}) = g_r\,(g_s - \delta_{rs})\,P\{\,Y_{rs\,k1,l1\,k2,l2}=1\} \qquad (5.10)$$

where g_r and g_s are the number of elements in groups G_r and G_s, and δ_{rs} is the Kronecker delta function (equal to1 if $r=s$, and 0 otherwise).

Hence, we can rewrite model (5.7) as follows:

$$\begin{aligned} \ln E\,(W_{rs\,k_1l_1\,k_2l_2}) = {} & \lambda_{rs} + (k_1 + l_1)\theta_1 + k_1\alpha_{1r} + l_1\beta_{1r} + l_1\alpha_{1s} + k_1\beta_{1s} + (k_1 l_1)\rho_1 + \\ & (k_2 + l_2)\theta_2 + k_2\alpha_{2r} + l_2\beta_{2r} + l_2\alpha_{2s} + k_2\beta_{2s} + (k_2 l_2)\rho_2 \end{aligned} \qquad (5.11)$$

For this model, we are estimating 32 parameters: 7 αs, 7 βs, 1 θ and 1 ρ for each relation. Since the unit of analysis of the model specified by (5.11) is still the dyad, there are 54(54-1) - 32 = 2830 degrees of freedom associated to the model. There are two other implications associated to having the dyad as the unit of analysis. First, the model specified by (5.11) still assumes independent dyads. Second, the likelihood-ratio goodness-of-fit statistic computed by commercial statistical packages is not correct, and it has to be calculated as specified by Fienberg and Wasserman (1981):

$$G^2 = -2\left(\sum_{r<s} \sum_{k,l} w_{rs\,k_1l_1\,k_2l_2} \ln(\hat{w}_{rs\,k_1l_1\,k_2l_2} / [g_r \cdot g_s\,]) + \sum_{r} \sum_{k,l} w_{rr\,k_1l_1\,k_2l_2} \ln(\hat{w}_{rr\,k_1l_1\,k_2l_2} / [g_r \cdot (g_r - 1)]) \right) \qquad (5.12)$$

Based on the assumption of equal parameters within subsets, our model has been greatly simplified. Indeed, we need to estimate 32 parameters for model (5.11) rather than 110 parameters for the original model (5.7). Furthermore, the standard chi-square distributions are more appropriate as reference distributions for the resulting test statistics. This is because the number of parameters is fixed and does not increase in the limit, as $g \rightarrow \infty$ (Haberman, 1981).

5.4.3. Testing the correlation between design interfaces and team interactions

The base model specified in equation 5.11 assumes that design interfaces and team interactions are two independent relations of the same network of elements. We will consider second-order interaction effects between design interfaces and team relations to capture the correlation between the design interface matrix and the team interaction matrix. We adapt the description of these effects provided by Wasserman and Iacobucci (1988) to our context as follows:

θ_{12} = parameter measuring tendency toward conformity across relationships. That is, component *i* depends on component *j*, AND team *i* reports interaction with team *j*.

ρ_{12} = parameter measuring tendency toward flow reversal. That is, component *i* depends upon component *j*, AND team *j* reports interaction with team *i*.

The θ_{12} parameter reflects the overall tendency toward positively associated design interfaces and team interactions (the "#" and the "blank" cells of Figure 3.X), and we expect this parameter to be statistically significant positive. On the other hand, the ρ_{12} parameter reflects the overall tendency toward dominance, that is, how likely it is that component *i* depending on component *j*, influences team *j* to interact with team *i*, and we have no reasons to expect this parameter to be statistically significant.

After extending the model with the second-order interaction parameters described above, the model can be written as follows:

$$\begin{aligned} ln\, E(W_{rs\,k_1l_1\,k_2l_2}) = {} & \lambda_{rs} + (k_1 + l_1)\theta_1 + k_1\alpha_{1r} + l_1\beta_{1r} + l_1\alpha_{1s} + k_1\beta_{1s} + (k_1l_1)\rho_1 + \\ & (k_2 + l_2)\theta_2 + k_2\alpha_{2r} + l_2\beta_{2r} + l_2\alpha_{2s} + k_2\beta_{2s} + (k_2l_2)\rho_2 + \theta_{12} + \rho_{12} \end{aligned} \qquad (5.13)$$

Holland and Leinhardt (1981) emphasize that testing whether $\theta_{12} = \rho_{12} = 0$ is the natural test to investigate for the existence of a true correlation between two matrices where each follows the p_1 distribution. Therefore, this base model provides a formal method to test the correlation of two design structure matrices that follow the p_1 distribution.

5.4.4. Fitting the base model to data

Wasserman and Iacobucci (1988) fit models of the form specified in (5.13) to analyze relation between actors at different points in time. Following the same rationale described previously on fitting p_1 models to data, the log-linear model for **W**-array that corresponds to the model (5.13) is

[12][13][24][14][23][15][26][16][25][34][56][35][46][36][45] (5.13)

Where the margins [35][46] allow for the inclusion of θ_{12}, while the margins [36][45] allow for the inclusion of ρ_{12}.

Testing the significance of any of the parameters in the model is relatively straight-forward using conventional rules for likelihood-ratio hypothesis tests for log-linear models for categorical data (for more details refer to Bishop, Fienberg, and Holland (1975)). Statistically, we must fit two models: 1) the null hypothesis model, with parameter restrictions given by H_o, and 2) the alternative hypothesis model, without the parameter restrictions. The difference in G^2 statistics for the two models then provides a test statistic for H_o, which is asymptotically $\chi 2$ with degrees of freedom equal to the difference in degrees of freedom for the two models.

Table 5.2 shows the results of fitting various models to data. Model 1 fits the independent model as specified in (5.11), and it serves as our reference model to test the null hypothesis ($\theta_{12} = \rho_{12} = 0$). The goodness-of-fit of model 1 indicates that the p_1 distribution provides a good fit to our data ($G^2 = 5243$, $df = 5692$). Model 2 improves the goodness-of-fit of model 1 by adding the parameter θ_{12} ($\Delta G^2 = 943.66$, $\Delta df = 1$, $p < .001$), while model 3 including ρ_{12} improves the goodness-of-fit of model 1 ($\Delta G^2 = 690.31$, $\Delta df = 1$, $p < .001$) but it is not as good as model 2. Finally, model 4 includes both second order parameters resulting in a model that is not statistically significant better than model 2 ($\Delta G^2 = 0.66$, $\Delta df = 1$, $p = .42$), which indicates than θ_{12} is statistically significant positive whereas ρ_{12} is not statistically significant.

It is important to note that the subscript 1 corresponds to the parameters of the design interface matrix, whereas the subscript 2 corresponds to the parameters of the team interaction matrix. The subscript FAN refers to the first group (r=1, s=1) while, the subscript LPC corresponds to the second group (r=2, s=2), and so on.

Table 5-2. Results of fitting base model to data

Parameters	Model 1	Model 2	Model 3	Model 4
		Parameters for design interface matrix		
α_{1FAN}	0.4321	**0.3647**	0.5379	0.3632
α_{1LPC}	0.2090	**0.3687**	0.3445	0.3632
α_{1HPC}	-0.0119	**0.1134**	0.0988	0.1061
α_{1BD}	-0.0171	**-0.2261**	-0.1974	-0.2137
α_{1HPT}	-0.5652	**-0.5774**	-0.6754	-0.5731
α_{1LPT}	-0.0770	**-0.0474**	0.0541	-0.0533
α_{1MC}	-0.2566	**-0.1597**	-0.3627	-0.1496
α_{1EC}	0.2869	**0.1638**	0.2001	0.1575
β_{1FAN}	-0.7415	**-0.6637**	-0.8206	-0.6587
β_{1LPC}	-0.0671	**0.1406**	0.0272	0.1355
β_{1HPC}	0.0509	**0.2175**	0.1288	0.2169
β_{1BD}	-0.0956	**-0.4275**	-0.2303	-0.4262
β_{1HPT}	0.4363	**0.3355**	0.4792	0.3360
β_{1LPT}	-0.3861	**-0.2278**	-0.3494	-0.2229
β_{1MC}	0.3176	**0.1866**	0.3883	0.1773
β_{1EC}	0.4856	**0.4384**	0.3766	0.4420
θ_1	-1.0653	**0.1633**	-0.1802	0.1032
ρ_1	3.9891	**3.3992**	3.5645	3.5224
		Parameters for team interaction matrix		
α_{2FAN}	0.2778	**0.24615**	0.4805	0.2312
α_{2LPC}	0.0061	**-0.2359**	-0.0917	-0.2327
α_{2HPC}	-0.0313	**-0.1505**	-0.1086	-0.1437
α_{2BD}	0.0008	**0.2441**	0.1583	0.2323
α_{2HPT}	-0.3079	**-0.0763**	-0.3376	-0.0678
α_{2LPT}	-0.0197	**0.0719**	0.1177	0.0706
α_{2MC}	-0.3880	**-0.3505**	-0.4613	-0.3491
α_{2EC}	0.4619	**0.2511**	0.2427	0.2595
β_{2FAN}	-0.5182	**-0.2435**	-0.5361	-0.2337
β_{2LPC}	-0.1838	**-0.3645**	-0.3424	-0.3553
β_{2HPC}	-0.1618	**-0.3141**	-0.2362	-0.3120
β_{2BD}	0.2766	**0.5727**	0.4218	0.5694
β_{2HPT}	0.3068	**0.2731**	0.4691	0.2591
β_{2LPT}	-0.5070	**-0.3641**	-0.4339	-0.3638
β_{2MC}	0.4624	**0.4074**	0.5224	0.4057
β_{2EC}	0.3248	**0.0327**	0.1351	0.0307
θ_2	-1.0619	**-0.1969**	-0.3954	-0.6043
ρ_2	3.5191	**2.3742**	2.7958	3.2314
		Second-order interaction parameters		
θ_{12}		**3.107**		3.2314
ρ_{12}			2.3158	-0.2101
		Goodness-of-fit		
G^2	5242.96	**4299.30**	4552.65	4298.64
DF	5692	**5691**	5691	5690

5.5. Including the Effects due to Organizational/System Boundaries

Before defining structural parameters that explicitly take into account the effects due to organizational/system boundaries, let us re-write hypotheses H3.1 and H3.6 as follows:

H5.1 Design interfaces are less likely to be matched by team interactions when they are across (organizational/system) boundaries. More specifically,

H5.1a When considering the dyads with design interfaces only, design interfaces across organizational boundaries are less likely to be matched by team interactions than the ones within boundaries (the "X" cells of Figure 3.1).

H5.1b When considering the dyads with team interactions only, team interacions across system boundaries are less likely to be predicted by design interfaces than the ones within boundaries (the "O" cells of Figure 3.1).

To explicitly represent organizational/system boundary effects, we define the following indicator variable:

ACROSS = 1 if elements (i.e. component and team) i and j are in the different groups ($r \neq s$)

ACROSS = 0 if $r=s$

By expanding the dimension of the **W**-array with ACROSS as the seventh dimension, we can estimate the parameter associated to the interaction ACROSS x k_1 x k_2 (due to symmetry of the **W**-array identical to ACROSS x l_1 x l_2). As shown in our base model the interaction terms k_1 x k_2 = l_1 x l_2 captures whether or not design interfaces are matched by team interactions. Hence, the third-order interaction effect that defines $\theta_{ACROSS,k1,k2} = \theta_{ACROSS,l1,l2}$ captures whether dyads across boundaries with design interfaces matched by team interactions are statistically significant different than dyads within boundaries with design interfaces matched by team interactions. Hence, a formal hypothesis testing of H5.1 can be specified as follows:

H5.1: $\theta_{ACROSS,k1,k2} < 0$

Since our model cannot differentiate the dyads with a state of either design interface with no team interaction ($[k_1=1]$ x $[k_2=0]$ = $[l_1=1]$ x $[l_2=0]$) or team interaction with no design interface ($[k_1=0] \times [k_2=1] = [l_1=0] \times [l_2=1]$), it is not possible to capture the hypothesized effects described in H5.1a and H5.1b. Therefore we are limited to formally test H5.1 only.

It is important to note that ACROSS is just an indicator variable and does not increase the number of states of the dyad. ACROSS it is completely defined by the independent states r and s. We consider this issue when fitting the data by defining structural zeros[6] (when using SPSS) for the dyads where ACROSS = 1 and $r=s$, and for the dyads ACROSS = 0 and $r \neq s$.

5.5.1.1. Testing Hypothesis *H5.1*

Table 5.3 shows the structural parameters of fitting the extended model to data. The BASE model corresponds with Model 2 from Table 5.2 (no structural parameters added). Model 1, Model 2 and Model 3 include second-order interaction effects which result in statistically significant negative effects indicating that fewer design interfaces and fewer team interactions take place across boundaries. Model 4 includes the third-order interaction effect of interest, which result (as expected) in a statistically significant negative parameter ($\Delta G^2 = 6.74$, $\Delta df = 1$, $p < .01$, supporting H5.1). Following conditional likelihood ratio tests (G^2 difference tests) Model 1 and Model 2 have been tested against the Base model, Model 3 has been tested against Model 2, and Model 4 has been tested against Model 3.

Table 5-3. Models extended to capture effects due to organizational/system boundaries

Structural Parameters	BASE	Model1	Model2	Model3	Model4
Second order interactions effects with ACROSS					
$\theta_{ACROSS,k1} = \theta_{ACROSS,l1}$		-1.644***		-0.8046***	-0.1595***
$\theta_{ACROSS,k2} = \theta_{ACROSS,l2}$			-1.9335***	-1.3042***	-1.0191***
Third order interaction effects with ACROSS					
$\theta_{ACROSS1,k1,k2} = \theta_{ACROSS1,l1,l2}$					**-0.9450*** **
G^2	4299.30	3189.26	3093.78	3068.3	**3061.56**
DF	5691	5690	5690	5689	**5688**

*<0.1 **<0.05 ***<0.01

Significance estimated with conditional likelihood ratio tests (G^2 difference tests) with 1 *df*

[6] Structural zeros are zeros imposed by design. Those are "zero occurrence" because such cases cannot occur by definition. For example WITHIN = 1 for a dyad whose elements belong to different groups.

5.6. Including the Effects due to Systems Modularity

In this section we test how the effects due to organizational/system boundaries are moderated by systems modularity. More specifically, we want to test the following hypothesis:

H5.2 The effects due to organizational/system boundaries are statistically significant different for interactions between modular systems than for interactions with integrative systems.

Similar to the previous section, we define structural parameters to include the effects due to systems modularity into the model. For that purpose, we define the following indicator variable:

MODULAR=1 if both components of a dyad belong to modular systems (r<7 and s<7)

MODULAR=0 if one of the components of a dyad belongs to integrative systems (r≥7 or s≥7).

By expanding the dimension of the **W**-array with MODULAR as the eighth dimension, we can estimate the parameter associated to the fourth order interaction effect MODULAR x ACROSS x k_1 x k_2 (due to symmetry of the **W**-array identical to MODULAR x ACROSS x l_1 x l_2). $\theta_{MODULAR.ACROSS.k1.k2}$ ($= \theta_{MODULAR.ACROSS.k1.k2}$) captures whether the effect due to organizational/system boundary is statistically significant different for modular systems than for integrative systems. To formally test hypothesis H5.2 the following condition needs to hold:

H5.2: $\theta_{MODULAR.ACROSS.k1.k2} \neq 0$.

Indeed, we expect this parameter to be less than 0, which corresponds with less cross-boundary design interfaces (matched by team interactions) between modular systems than with integrative systems.

It is important to note that MODULAR is just an indicator variable and does not increase the number of states of the dyad. MODULAR is completely defined by the independent states r and s. This issue is considered when fitting the data by defining structural zeros (when using SPSS) for the cases where MODULAR = 1 and ($r \geq 7$ or $s \geq 7$), and for the cases where MODULAR = 0 and ($r<7$ and $s<7$).

5.6.1.1. Testing Hypothesis *H5.2*

Table 5.4 shows the structural parameters of fitting the extended model to data. The ACROSS model correspond with Model 4 from Table 5.3. Model 1, Model 2 and Model 3 include second-order interaction effects with MODULAR which result in not statistically significant effects indicating that the distribution of design interfaces and team interaction for modular systems is not statistically significant different than for integrative systems. Model 4 includes the third-order interaction effect with MODULAR which also result in a not statistically significant parameter indicating that the distribution of design interfaces matched by team interactions for modular systems is not statistically significant difference than for integrative systems. Model 5 includes the four-order interaction effect of interest which resulted to be statistically significant negative supporting hypothesis H5.2.

Table 5-4. Models extended to capture moderating effects due to systems modularity

Structural Parameters	ACROSS	Model1	Model2	Model3	Model4	Model5
Second order interactions effects with ACROSS						
$\theta_{ACROSS,k1} = \theta_{ACROSS,l1}$	-0.1595***	-0.1028***	-0.9671***	-0.1029***	-0.1162***	**-0.5517*****
$\theta_{ACROSS,k2} = \theta_{ACROSS,l2}$	-1.0191***	-1.0157***	-0.154***	-1.0153***	-1.0207***	**-1.469*****
Third order interaction effects with ACROSS						
$\theta_{ACROSS,k1,k2} = \theta_{ACROSS,l1,l2}$	-0.9450***	-0.9512***	-0.9569***	-0.9512***	-0.936***	**-0.6188*****
Second order interaction effects with MODULAR						
$\theta_{MODULAR,k2} = \theta_{MODULAR,l2}$		0.2187		0.2180	0.1697	**-0.0961**
$\theta_{MODULAR,k2} = \theta_{MODULAR,l2}$			0.1725	0.0014	-0.0228	**0.0444**
Third order interaction effects with MODULAR						
$\theta_{MODULAR,k1,k2} = \theta_{MODULAR,l1,l2}$					0.0571	**0.4768**
Four order interaction effects with ACROSS and MODULAR						
$\theta_{MODULAR,ACROSS,k1,k2} = \theta_{MODULAR,ACROSS,l1,l2}$						**-0.3354*****
G^2	3061.56	3060.83	3060.79	3060.78	3060.76	**3030.81**
DF	5688	5687	5686	5685	5684	**5683**

*<0.1 **<0.05 ***<0.01

Significance estimated with conditional likelihood ratio tests (G^2 difference tests) with 1 *df.*

5.7. Discussion of the Results

Based on previous research from social science we acknowledge that team interaction data reported by design team members are likely to present effects of reciprocation, differential expansiveness and differential attraction. Similarly, we assume that design interface data reported by design experts are also subject to these effects. These observations are the main threats to validity of the results presented in chapter four, and the main motivation of the analysis presented in this chapter.

We built upon research completed in the field of social network analysis to develop a base log-linear model that allows for the simultaneous estimation of parameters that measure the average tendency toward reciprocation (ρ), and the amount of differential expansiveness and differential attractiveness of each element of the network (α's and β's, respectively). We then extended the base model by adding structural parameters that capture the effects of organizational/system boundaries and the effects of system modularity. We confirmed the results that design interfaces are more likely to be matched by team interactions when they take place within organizational/system boundaries. Also, we found that the effect due to organizational/system boundaries is moderated by systems modularity.

Three major limitations of the model presented in this chapter are the following:

- It assumes statistically independent dyads, therefore we cannot use this model to study triadic effects. Triadic effects are associated to the hypothesized effects due to potential indirect team interactions and to potential secondary dependencies. Future research might take advantage of more recently developed models (Wasserman and Pattison, 1996) that better handle triadic effects.
- Structural parameters have to be defined for all the dyads in the network. For example, the effects due to design interface strength cannot be tested with this model because it is not possible to define a binary indicator variable "STRENGTH" for the dyads that do not share an interface. That is, there would not be any difference between a dyad that shares a weak design interface (STRENGTH=0) and a dyad that does not share a design interface (STRENGTH=0).
- Even though commercial statistical software packages can be used to fit the model to empirical data, the data obtained from the printouts needs further manipulation to estimate the model's parameters (see Appendix D). Also, the G^2 obtained from the statistical packages is wrong and needs to be corrected.

6. *Geographically Distributed Product Development: An Empirical Study in the Telecommunications Industry*

La Pie, 1868-69
Claude Monet

6.1. Introduction

The dynamics of current businesses have challenged the execution of product development projects by increasingly requiring more geographically distributed teams to work together (Chen and Bolon, 1993; De Meyer, 1993; Granstrand *et al.*, 1993; Griffin, 1997). Current practices in product development involve the execution of various stages of the process in various locations around the globe. It is common to encounter firms that design their hardware in one location, their software in another location, while having their manufacturing facilities yet spread to other locations. Ghoshal *et al* (1990) recognize the importance of developing products in a distributed fashion when serving diverse markets. McDonough *et al* (1999) present the challenges associated with managing global new product development. Leonard *et al* (1998) present a case study of a geographically distributed software development project, illuminating the problems faced when managing these types of virtual organizations.

Many researchers have also recognized the tremendous changes occurring in the way current organizations communicate (Yates and Orlikowski, 1992). The use of electronic-based communication media is increasing the number of options distributed development teams have available to coordinate activities, to keep knowledge up-to-date and to spark creativity with non-collocated team members. The widespread use of information technology is reducing the traditional reliance on face-to-face communication in what has been called the "network organization" (Sproull and Kiesler, 1991).

While previous research demonstrating the negative relation between communication and distance is well established (cf. Wells, 1965; Conrath, 1973; Allen, 1977; Keller and Holland, 1983; Pinto *et al.*, 1993), less is known about how the relationship varies with different types of media or communication content nor how distance affects the choice of media used (Van den Bulte and Moenaert, 1998). Utilizing a rich empirical data set collected from interviews in three geographically distributed development teams in the telecommunications industry, we analyze the moderating effects of communication media and content on the relation between communication frequency and distance. In addition, we examine how distance, and other moderating variables, affects the choice of communication media.

6.2. Related Work in Communication in Global Product Development

Ghoshal and Bartlett (1988) reported findings from an empirical study of sixty-six North American and European multinationals indicating that subsidiaries with higher levels of inter-unit communication were more effective in the creation, adoption and diffusion of innovations. In their study of global new product development teams (GNPDT's), McDonough *et al.* (1999) correlated GNPDT performance with the use of multiple communication mechanisms, what they called an "affiliated set," consisting of phone, fax, email, teleconferencing and company databases.

While communication patterns in product development depend on the nature of the project and the organizational structure executing it (Barczac and Wileman (1991); Morelli *et al.* (1995)), distance also plays an important role (Allen, (1977); De Meyer (1991)). The negative influence of distance on communication has been studied so extensively as to be "accepted as an axiom in social theory" (Van den Bulte and Moenaert (1998: p. S3)). Allen's (1977) research on the communication processes in R&D organizations, describing how increasing distance between team members reduced the chances of two team members communicating for technical matters, is probably the best known of these studies in the R&D context. However, there have been several, more recent studies supporting his general findings (cf. Keller and Holland (1983); De Meyer and Mizushima (1989); Jaffe *et al.* (1993); Pinto *et al.* (1993); Van den Bulte and Moenaert (1998)).

Taking exception with much of the previous research on the influence of distance on communication, Van den Bulte and Moenaert (1998) claim that "previous research does not allow one to conclude confidently that distance is a major barrier to communication in R&D settings" (p. S3). They note that much of this research lacks contextual realism, internal validity and statistical conclusion validity. Utilizing statistical modeling techniques for sequential network data (Wasserman and Iacobucci (1988)), Van den Bulte and Moenaert examine a "naturally occurring managerial intervention involving the relocation of R&D teams in a leading high-tech company" (1998: p. S4). Although they found that collocation of R&D team members did enhance communication among the members of the team, they also discovered that the communication frequency between R&D and marketing was not affected by the resulting increase in physical distance. "This unexpected asymmetric result suggests that the effect of distance on communication may be moderated by the nature of the communication. Because we measured oral communication broadly, *without discriminating between various media or contents*, directly testing such a conjecture must be left for future research" (Van den Bulte and Moenaert (1998: p. S15), emphasis added).

6.3. Formulating the Hypotheses

Similar to Hightower and Sayeed's (1996) "opportunity" and "motivation," we divide the factors that influence technical communication into two categories: *communication drivers* and *communication barriers*. We define communication drivers as the factors that motivate information transfer between interacting team members, and communication barriers as the factors that hinder the process of exchanging information.

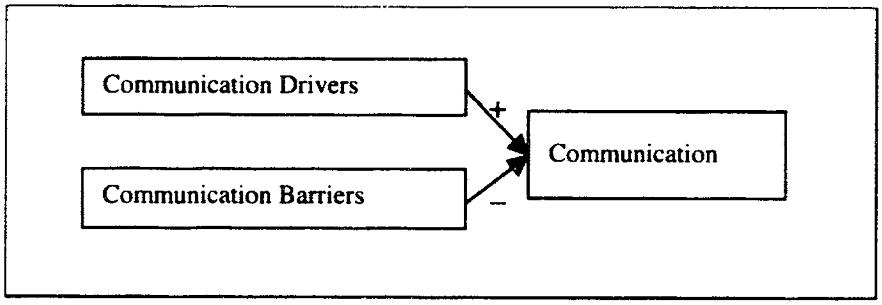

Figure 6-1. Factors that influence technical communication

6.3.1. Communication Drivers

In the organizational communication literature, Daft and Lengel (1986) present an integrated framework, based on the concepts of *uncertainty*—the absence of critical and stable information and *equivocality*—the lack of understanding of a situation—to explain what drives information processing in organizations[7]. Similarly, technical communication in product development is required to reduce information deficit—that is, team members deal with unstable information and so must communicate critical parameters as they become known—and to reduce ambiguity—that is, team members deal with imprecise information and so must communicate to define problems or to reach consensus on the solution of a problem. This is similar to the concepts of coordinative information—that used simply to coordinate activities—and innovative information—that used in problem solving—described by Hauptman (1986).

The degree of task *interdependence* describes the degree to which tasks require collective action (Wageman, 1995): The greater the degree of task interdependence, the greater the coordinative and innovative information requirements (De Meyer, 1991). This is consistent with previous research that has shown that a greater degree of task interdependence leads to greater communication (Crawford and Haaland, 1972; Adler, 1995; Allen, 1997). Allen (1997) recognizes that the degree of interdependence between engineers' work is directly related to the probability that they engage in frequent technical communication. At the task level, Smith and Eppinger (1997) use the strength of task interdependency to identify the activities that require higher effort to coordinate. Loch and Terwiesch (1998) present an analytical model to study the

[7] They were primarily concerned with managerial communication instead of technical communication.

coupling of uncertainty, dependence and communication, suggesting that average communication frequency increases with the level of uncertainty and dependence[8]. These results are consistent with the empirical evidence presented by Adler (1995) and the numerical approach presented by Ha and Porteus (1995).

Thus, we propose the following hypothesis regarding the effects of interdependence on communication frequency:

H6.1: Communication frequency increases with the degree of interdependence, independently of the communication media used.

Although the majority of technical communication among interacting team members is likely to involve coordinative and innovative information, these are not the only types of communication. Team members may also engage in technical communication for inspiration and general knowledge, not directly related to specific development tasks (Morelli *et al.*, 1995; Allen, 1997). Team members can communicate for creative inspiration, managerial affirmation, and to keep up to date with the latest developments in their disciplines. In addition, there is a general tendency for individuals to seek out similar others with which to communicate—what Van den Bulte and Moenaert (1998) refer to as homophily effects.

Organizational structure establishes boundaries within the organization. People within such boundaries are subjected to organizational bonds that promote the development of a language and an identity inherent to the group. Allen (1977) found that organizational bonds[9] increased the probability of two team members engaging in technical communication. Thus, we expect the following hypothesis to hold true:

H6.2: Communication frequency is higher between individuals who share an organizational bond, independently of the communication media used.

---

[8] They model uncertainty as the number of design changes, and dependence similar to the "downstream sensitivity" of Krishnan et al (1997).

[9] Other terms are: organizational affiliation, organizational ties.

6.3.2. Communication Barriers

There are several factors opposing technical communication between members of a product development team. The literature suggests three major types of geographic barriers to the communication process:

- physical distance
- overlapping working time, and
- cultural/language differences.

As stated above, there is considerable empirical research demonstrating the negative effects of distance on technical communication. Allen summarizes his findings about how individual location influences technical communication in the "communication-distance" curve for face-to-face communication in collocated R&D organizations (Allen (1977, p. 239)). Allen found that the probability of two engineers engaging in technical communication rapidly decays with distance, and suggested that such a communication pattern is *independent* of the medium used to communicate (Allen (1997)). It is important to note that Allen's (1977, 1997) results imply that distance is a discontinuous function, that is, "it is only within the first thirty meters that separation has any real effect on the probability of communication" (Allen (1977, p. 240)). Allen's work uses distance as a proxy for a wider issue of the influence of architecture on communication. On the other hand, we use distance to capture separation from a global point of view.

Although it is not difficult to hypothesize how physical distance presents a direct barrier to face-to-face communication, it is less clear why physical distance would reduce communication independently of the media used. One possible explanation is the concept of the "affiliated set" of communication mechanisms that support each other (McDonough *et al.* (1999)). De Meyer (1991) found in his studies of global R&D that "(o)ther than calls for simple exchanges of data, one only calls the people one knows well and sees fairly often." Thus, one might expect a positive correlation in the communication frequency among various media. As distance reduces face-to-face communication, there is a correlated reduction in the use of other media.

H6.3: Communication frequency decreases with distance, independently of the communication media used.

In addition, as distance increases, so might working-time differences. With decreasing overlapping working time, synchronous communication would become more difficult. Under the hypothesis that communication frequency is correlated among the various media, then asynchronous communication would also decrease.

H6.4: Communication frequency increases with overlapping working-time, independently of the communication media used.

Another possible explanation is that distance is a proxy for other factors such as culture, language and identity. With increasing distance, one would expect increasing differences in language and cultural identity, and thus homophily—one of the proposed drivers of communication—independent of any organizational bonds that might be shared. Thus, if one could measure differences in language and culture directly, one could identify the effects of these on communication.

H6.5: Communication frequency decreases with cultural/language differences, independently of the communication media used.

6.3.3. Media Choice:

Before formulating the effects of geographic dispersion on media choice, we map three communication media (face-to-face, telephone and email), commonly used in organizations, in a space-time domain (Figure 6.2).

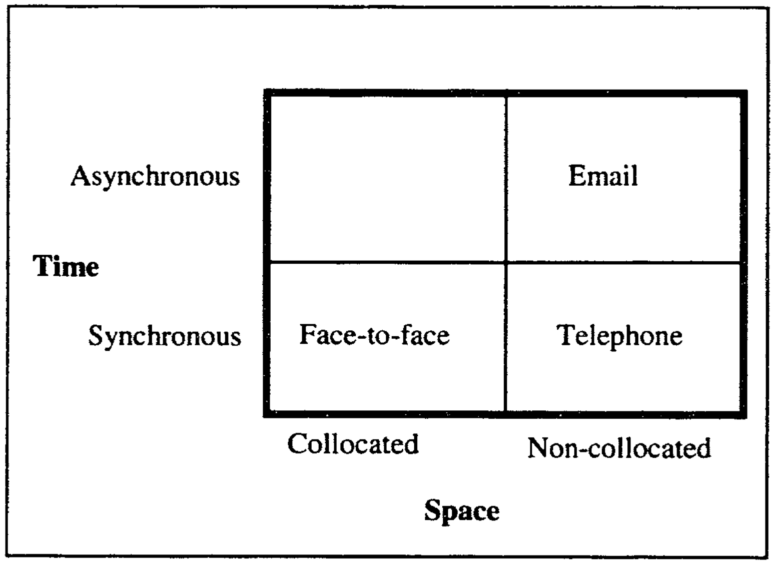

Figure 6-2. Time-space domain of communication media.

Although H3 predicts that distance reduces communication frequency across all media, one would expect that the magnitude of the impact would differ. Given the need of physical proximity for face-to-face communication, we would expect distance to have a much greater impact on face-to-face communication than for non-collocated communication.

H6.6: The rate of decay depends upon the communication medium used. Face-to-face communication would exhibit faster decay than non-collocated communication such as telephone and email.

This, of course, has an implication for the choice of media used. Media richness theory (Daft and Lengel, 1984)—one of the most broadly studied theories about media choice—ranks communication media according to their capacity to process ambiguous information. Specifically, they rank media based upon their ability to provide *feedback*, their capacity to transmit *multiple cues*, their availability to use *natural language,* and their *personal focus*. According to the theory, face-to-face is a richer medium than telephone, and telephone is a richer medium than email. This theory provides a rational criterion to select media to reduce ambiguity.

Empirical evidence has supported this theory for managerial type communications (Trevino *et al*, 1987, Jones *et al* (1988), Schmitz and Fulk (1991)).

Allen and Hauptman (1990) agree with media richness ranking when comparing the bandwidth of certain communication media. However, they also rank media according to their *data-transmission efficiency*. They argue that email is a more efficient medium than telephone and face-to-face from a data-transmission standpoint[10]. By ranking communication media by a data-transmission efficiency standpoint they provide a rational criterion to select media to reduce information deficit. This criterion is particularly relevant when large amounts of information, such as CAD models, analysis results, and design or manufacturing specifications, need to be transferred.

While "improvements in information technologies will make it easier for technical professionals to communicate, ... knowledge is best transferred to engineers through personal contact" (Allen and Hauptman, 1990: p. 282-284). Other authors have addressed the issue of effectiveness and efficiency of communication media (Hiltz and Turoff, 1978; Siegel *et al*, 1984). Warkentin, et al. (1997) found that although virtual and face-to-face team interactions exhibited similar levels of communication effectiveness, teams using face-to-face interactions reported higher levels of satisfaction with team performance.

Given the fact that face-to-face is a synchronous, collocated medium (see Figure 6.2) we expect its probability of being used to rapidly decay with distance, whereas the probability of using an asynchronous, non-collocated medium such as email should grow with distance. When product development teams are distributed around the globe, effects of distance are compounded by the time zone difference between the interacting team members. Its major effect is that simultaneous working time reduces, increasing the efforts to have synchronous communication or simply fast feedback (Gulati and Eppinger (1996), McDonough *et al* (1999)). Telephone (a synchronous, non-collocated medium) may be preferred for distant communication as long as there is simultaneous working time (low time zone difference). Finally, email (an asynchronous, non-collocated medium) will be preferred for long-distance communication. As a result, we formulate the following hypotheses:

---

[10] Marril (1980, p. 185) discusses in more detail the efficiency of transmitting digital data.

H6.7a) The probability of using face-to-face communications rapidly decays with distance.

H6.7b) The probability of using telephone communication increases, reaches a maximum and then decays with distance.

H6.7c) The probability of using email communication increases with distance.

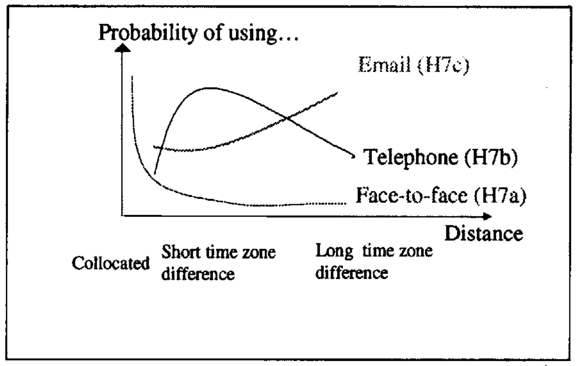

Figure 6-3. Effects of distance on media choice (H6.7)

Geographical separation also implies, in many cases, cultural difference. Language differences, different customs, different ways of referring or treating others have all been recognized as a major barriers to communication (McDonough III and Kahn, 1996; Gulati and Eppinger, 1996; Leonard *et al.*, 1998). Language differences, in particular, create the need for written-asynchronous communication, which allows interacting parties to take more time to interpret and process the information exchanged (McDonough *et al.*, 1999). Thus, we formulate the following hypothesis regarding the influence of language difference on media choice:

H6.8: The probability of using written-asynchronous communication media, such as email, rather than verbal-synchronous communication media, such as telephone, increases with language difference.

6.4. The Empirical Study

In the spring and summer of 1995, more than 200 interviews were conducted at 30 facilities, in 13 countries, in three large multinational corporations (MNCs) in the telecommunications industry. These interviews were part of a three-year study of innovation in MNCs conducted from 1994 to 1996. The interviews, which lasted anywhere from one to three hours, were structured with a list of questions, taped and later transcribed. In addition, field notes were taken, and forms were filled out when quantitative data were requested. The transcriptions, field notes and data forms were then used to construct a systematic data set.

Communication data were collected by interviewing members of three different development teams at three different companies within the telecommunications industry. In each interview the respondent was asked to give the name, location and position (including functional affiliation) of the people he/she communicated with during the project. Respondents were asked repeatedly to give us as complete a list as possible. However, on a few occasions, respondents would tire after about two pages (20 partners).

For each communication partner, respondents were asked to rank from 1 (lowest) to 10 (highest) the importance of the communication for the execution of their project-related tasks. Additionally, the medium selected to communicate as well as the communication frequency was reported per each interaction. Finally, a brief, qualitative description of the content of the communication was requested.

Given information on the location of the respondent and their communication partners, we estimated the *distance* (in kilometers) between the facilities where individuals were located. Communication partners located at the same facility were given a distance of zero, regardless of the particular "micro-location" of their offices. We also determined the time zone difference to calculate the *overlapping working time. Language differences* were estimated based on the location and name of the respondent and communicating partners. Examining the title, position description and role in the project of the respondents and their communication partners we determined the level of their *organizational bonds* (either function organizational bonds or project organizational bonds). Since the general content of the message exchanged was also provided for each interacting pair, we grossly estimated the *type of technical communication* associated to each interaction. A detailed description of the variables used in our analysis is provided in Table 6.1.

Table 6-1. Description of variables used in the analysis.

Metric	Description
Importance of the interaction	Scale metric that measures the level of criticality of the interaction from the respondent standpoint. It assesses the degree of task interdependence associated to each interacting pair. A scale from 1 to 10 was used (1=low importance, 10=high importance).
Organizational bonds	Binary metric to capture the level of organizational affiliation between interacting parties. 0=weak organizational bond such as different organizations, different tasks, different professional background. 1=strong organizational bond such as same organization, similar tasks, similar professional background.
Distance	Distance (in kms) between the cities where each of the parties was located.
Overlapping working time	Number of hours in which both parties would be in their office simultaneously (assuming working hours to be from 9 am to 5 pm).
Language difference	Binary variable. 0=same native language. 1=different native language.
Communication frequency using certain communication medium	Number of interactions per week using certain communication medium (face-to-face, telephone and email).

Some researchers (Tushman, 1978; Allen and Cohen, 1969) have already attempted to measure information processing by counting communication transactions such as number of memos, number of telephone conversations or face-to-face communications. We also use communication frequency (i.e. number of interactions per unit time per each communication medium used) as our dependent variable.

However, it is important to note that our metrics for capturing technical communication differ from the ones used by Allen (1977). Allen determined the probability of two researchers engaging in technical communication as a function of distance. Allen determines such probabilities by dividing the number of team members who communicate (at least once a week) by the total number of people available at each distance range. Allen considered all potential pairs in the development organization. Given the scale of our project, it is impractical for us to use the same approach. Instead, we consider only the pairs that actually communicate and their absolute and relative use of communication media to exchange technical information, from the respondent's point of view.

From 255 interviews (respondents) we obtained a total of 829 interacting pairs (dyads) that formed the initial raw data. A screening to eliminate pairs with missing and/or inconsistent information reduced the data set to a sample of 653 interacting pairs of which 485 pairs contained complete information for all the variables. Table 6.2 shows the descriptive statistics of the sample data analyzed.

Tables 6.3 and 6.4 show correlations among the independent and dependent variables, respectively. As one might expect, overlapping working time is highly correlated with distance (-0.945). This will make it difficult to disentangle the two in the analysis. Language is also positively correlated with distance, and thus, with overlapping working time, but to a much lesser extent (0.599). Also as expected, the frequency of communication in face-to-face is positively correlated with the frequency of communication in telephone, but only slightly (0.245). Email, though positively correlated with face-to-face (0.172), has very little correlation with telephone (0.027).

Table 6-2. Descriptive statistics

	IMPORTANCE (1-10)	ORG. BONDS (0/1)	DISTANCE (kms)	OVERLAP TIME (hours)	LANGUAGE (0/1)	FACE-to-FACE FREQ (#/week)	TELEPHONE FREQ (#/week)	EMAIL FREQ (#/week)
Mean	6.94	0.476	1,922	6.68	0.222	1.325	0.574	0.935
Maximum	10.0	1.0	15,658	8.00	1.0	25.0	20.0	35.0
Minimum	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Std. Dev.	2.36	0.4999	3,754	2.48	0.416	2.279	1.584	3.105
Skewness	-0.569	0.0949	2.157	-1.708	1.333	4.061	6.934	7.663
Kurtosis	2.495	1.009	6.862	4.492	2.777	32.003	68.349	71.424
Jarque-Bera	31.36	80.83	677.6	280.7	144.7	18332.0	90188.7	99359.9
Observations	485	485	485	485	485	485	485	485

Table 6-3. Correlation between the independent variables

	IMPORTANCE	ORG. BONDS	DISTANCE	OVERLAP TIME
IMPORTANCE	**1.000**			
ORG. BONDS	0.101	**1.000**		
DISTANCE	-0.049	-0.088	**1.000**	
OVERLAP TIME	0.051	0.096	-0.945	**1.000**
LANGUAGE	-0.052	-0.007	0.599	-0.554

Table 6-4. Correlations between the dependent variables

	FACE-to-FACE FREQ	TELEPHONE FREQ
FACE-to-FACE FREQ	**1.000**	
TELEPHONE_FREQ	0.245	**1.000**
EMAIL_FREQ	0.172	0.027

6.5. Statistical Analysis

Several studies have posited and examined the effects of subsidiary type and MNC strategy on patterns of communication (c.f. Ghoshal and Bartlett, 1990; Nobel and Birkinshaw, 1998). Although we have not set out to examine explicitly these relationships, we must be aware that firm-level characteristics could have a significant influence on communication. The three multinationals that we studied were all in the telecommunications industry, but each was headquartered in a different country: Sweden, Japan and North America.[11]

The Swedish MNC was the furthest along in terms of the internationalization of its new product development. Most of its facilities could be classified as *international creators* in the typology of Nobel and Birkinshaw (1998). The Japanese MNC was the least internationalized, with many of its facilities evolving from *local adopters* to *international adopters*. The North American MNC was in between, but closer to the Swedish MNC in internationalization of new product development.

The three projects that we studied consisted mainly of software development—though with some, more or less related hardware developments as well. The projects in the Swedish and North American MNCs each involved the development of a global product platform. The project in the Japanese MNC involved the development (adaptation) of a product local to the North American market. Figure 6.4 plots the dyad-distance profiles for each of the three project samples. It highlights the difference of the Japanese distance profile from those of the other two MNCs. In light of this evidence, we ran separate analyses for each firm. However, because the results were not statistically significantly different from the pooled data, the results presented in this section are for the pooled data only.

---

[11] The "North American" corporation had its original headquarters and basic research labs in Canada, but had recently moved the headquarters for the particular business unit that we were examining to the US.

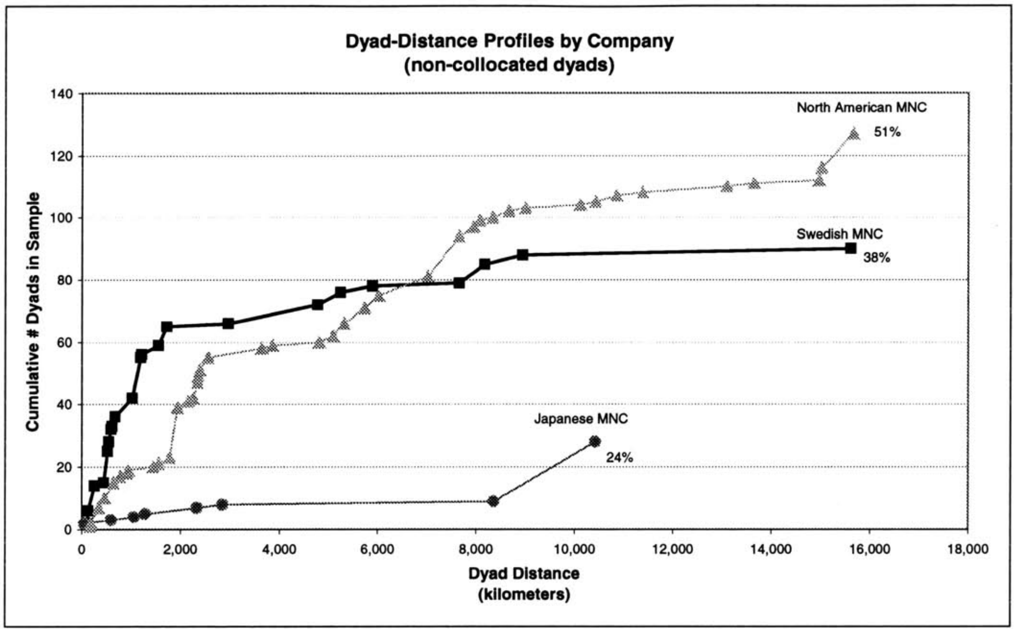

Figure 6-4. Dyad-distance profiles for each of the three project samples

6.5.1. Communication Frequency

We completed several linear regression models whose results are compiled in Table 6.5. The first column of this table contains the independent variables. The rest of the columns contain the non-standardized coefficients included in each of the models. An empty cell indicates that such a variable was excluded from the model due to lack of significance.

The dependent variable of the models exhibited in Table 6.5 is the natural log of communication frequency. This specification of the dependent variable has three important implications.

1. *ln*(communication frequency) is closer to a normal distribution, supporting the assumption that the errors of the regression models are normally distributed.
2. the negative coefficients of *ln*(distance+1.0) can be interpreted as the rate of decay of communication frequency due to distance.
3. the coefficients of the other variables included in the models provide an approximation of the percentile change in communication frequency given a unit change in the corresponding variable.

Table 6-5. Regression results for communication frequency†

Independent Variables	Model 1 (Total)	Model 2 (Face-to-face)	Model 2' (Face-to-face)	Model 3 (Telephone)	Model 4 (Email)
Constant	-0.916***	-2.463***	-1.239***	-1.505***	-1.495***
Importance	0.191***	0.199***	0.200***	0.161***	0.184***
Organizational bonds	0.402***	0.155***	0.341***	0.678***	0.653***
Distance					
ln(distance+1.0)	-0.117***	-0.199***	-0.254***	-0.075***	-0.064***
Overlapping working time		0.305**			
Language difference					
N	485	298	298	213	224
Adj. R^2	0.290	0.452	0.445	0.213	0.260

† $\text{communication frequency} = e^{(\alpha_0 + \alpha_1 \cdot \text{importance} + \alpha_2 \text{ organizational bonds})} \cdot (\text{distance} + 1.0)^{\alpha_3}$

* <0.1; ** <0.05; *** <0.01

The first model shown on Table 6.5 (Total) refers to *total communication frequency*, defined as the summation of all three communication frequencies (i.e. face-to-face, telephone and email communication frequencies) associated with each interacting pair. Models 2-4 are separate runs for each media type. The results clearly support hypotheses H6.1-H6.3. That is, communication frequency increases with the *importance of interaction* (H6.1) and with the presence of strong *organizational bonds* (H6.2), but decreases with *distance* (H6.3) across all media.

Not surprisingly, given its strong correlation with distance, the results for *overlapping working-time,* and thus, for hypothesis H6.4, are mixed. For face-to-face communication, overlapping working time is significant at the 0.05 level, but is not statistically significant for total communication or for the other two media. Given its correlation with distance, we exclude it in model 2'. The results do not support the homophily hypothesis (H6.5) that communication frequency decreases with *language differences* for any media.

By looking at the coefficients of *importance of the interaction* and their standard error for each of the models, and thus, for each media, we observe no statistically significant difference among them. Furthermore, importance of interaction explains about the same amount of variation for each of the models (media). As a result, we can conclude that the effect of importance of interaction is fairly consistent across all media used.

When analyzing the effects of *organizational bonds* on each of the models, we observe that both telephone and email communication frequencies are much more sensitive to the presence of strong organizational bonds. Additionally, organizational bonds explain a greater portion of variation of telephone and email communications than they do for face-to-face communications.

As we hypothesized (H6.6), the effect of *distance* on communication frequency is significantly contingent to the media used. For face-to-face communication, the rate of decay in communication frequency and the amount of variation in the data explained by *ln*(distance+1.0) is much greater than for telephone and email communications.

6.5.2. Media Choice

In order to explore the effects of degree of interdependence, organizational bonds and geographic dispersion on media choice, we derive a relative communication frequency per medium by dividing each communication frequency per medium by the total communication frequency associated to each interacting pair. That is, we define the probability that an interacting pair uses a certain communication medium as follows:

$$\text{P(using medium)} = \frac{\text{Communication frequency of medium}}{\text{Total communication frequency of all media}} \text{ per interacting pair.}$$

In order to test the effect of *distance* on media choice we ran linear regression models that include *distance* and *ln(distance+1.0)* as independent variables. The results are shown in Table 6.6 and the resultant curves that describe these models are graphed in Figure 6.5. The dependent variable of the models shown in Table 6.6 is the natural log of the probability of using either face-to-face, telephone or email, respectively. We included the *p*-values (between parentheses) of the variables included that were not significant.

Table 6-6. Results for the effects of distance on media choice

Independent Variables	P(face-to-face)	P(telephone)	P(email)
Constant	0.476***	0.138***	0.118***
Importance	-0.002 (0.696)	-0.001 (0.859)	0.005 (0.251)
Organizational bonds	0.034 (0.103)	-0.023 (0.293)	-0.027 (0.212)
Distance	6.96E-6 (0.113)	-1.88E-5***	9.97E-6**
ln(distance+1.0)	-0.055***	0.040***	0.016***
N	485	485	485
Adj. R^2	0.427	0.173	0.132

*<0.1; **<0.05; ***<0.01 (*p*-values within parentheses)

$$\text{probability of using certain medium} = e^{(\alpha_0 + \alpha_1 \cdot \text{distance})} \cdot (\text{distance} + 1.0)^{\alpha_2} - 1.0$$

The results presented in Table 6.6 and graphed in Figure 6.5 support hypotheses *H6.7a-c*. That is, the probability of using face-to-face rapidly decays with *distance,* the probability of using telephone increases, peaks and then decays with *distance*, while the probability of using email increases with *distance*. Given the significant and consistent influence of importance and organizational bonds on communication frequency across all media, it is interesting to note that neither importance nor the presence of organizational bonds is shown to influence media choice.

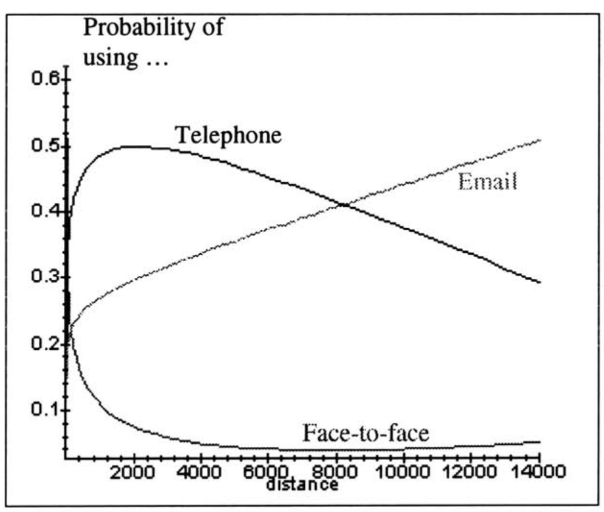

Figure 6-5.Distance-based linear regression results

As noted previously, distance can be a proxy for language and working time differences. Table 6.7 presents the new results when we add these later two variables to the model. Again, the results for *overlapping working time* are not significant. However, we see the *language differences* are significantly negatively correlated with the use of telephone and positively correlated with the use of email. The results in Table 6.7 support hypothesis *H6.8* that the probability of using written-asynchronous communication media, such as email, rather than verbal-synchronous communication media, such as telephone, increases with language difference.

Table 6-7.Results for media choice with language and working time

Independent Variables	P(face-to-face)	P(telephone)	P(email)
Constant	0.462***	0 122***	0.153***
Importance			
Organizational bonds	0.033 (0.110)		-0.035*
Distance		-1.66E-5***	2.15E-6 (0.224)
ln(distance+1.0)	-0.046***	0.054***	
Overlapping working time			
Language difference	-0.054 (0.117)	-0.199***	0.269***
N	485	485	485
Adj. R^2	0.429	0.222	0.239

*<0.1; **<0.05; ***<0.01. *p*-values in parentheses.

6.6. Discussion of Results

Even though we completed another study on how distance negatively influences communication in distributed development organizations, this study makes important contributions along many dimensions. First, noting Van den Bulte's and Moenaert's (1998) comments about "contextual validity", our study examined communication within three global new product development teams (GNPDTs). Interviews were conducted during the actual development project, and so did not rely on the ability of respondents to recall details of previous experiences. Second, our study is on a much different scale then many others—notably, Allen's (1977) often-cited study of collocated R&D personnel and Van den Bulte and Moenaert's (1998) study of the relocation of R&D personnel into another building. Our study is more on the macro scale of "global" dispersion in international development activities. Finally, and most importantly, we not only studied the use of different communication media, which allows us to discriminate the effects of geographic dispersion among various media, but also we found that the negative influence of distance can be compensated by high degree of team interdependence, strong organizational bonds, and use of electronic-based communication media. Given the empirical results presented in this paper a more sophisticated version of Figure 6.1 is exhibited in Figure 6.6.

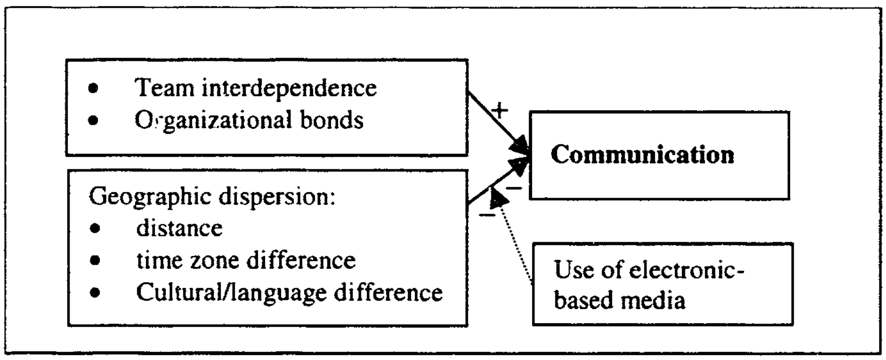

Figure 6-6. Summary of results

Consistent with previous research, we found that both *interdependence* (as measured by the importance of the interaction) and *organizational bonds* were positively correlated with communication frequency across all media. This observation supports the hypotheses that interaction criticality and homophily are major communication drivers. The surprising result was that neither of these two independent variables was correlated with media choice. Apparently, people involved in critically interdependent tasks or who share strong organizational bonds engage in broad spectrum of communication means.

Even when team members were non-collocated higher communication frequencies were observed for highly interdependent pairs. These results reinforce the importance for managers identifying critical tasks dependencies in their organizations in order to facilitate intense communication among the team members involved in such interdependent tasks. Furthermore, managers can overcome the negative effects of distance by constantly notify their team members about the level of criticality of their interdependence.

Conversely, by documenting communication frequencies managers can uncover the underlying structure of development projects as illustrated by McCord and Eppinger (1993). Since the effect of importance of the interaction on communication frequency is fairly consistent across all media used, we can track electronic-based communication transactions to easily identify team dependencies, especially when teams are geographically distributed. Tracking electronic-based communication frequencies can provide an easy and non-disruptive way to obtain the dependency-structure[12] of a development project.

Although we also found supporting evidence to the hypothesis that communication frequency increases in the presence of strong organizational bonds, the surprising finding was the moderating effect of media used. As evidenced by our results, strong organizational bonds have stronger positive effect on telephone and email communications than in face-to-face communications. Therefore, *organizational bond* is another element that can help managers to overcome the negative influence of distance on technical communication.

As hypothesized, *distance* between interacting pairs negatively correlates with communication frequency across all media. However the magnitude of this effect depends upon the medium used to communicate. Face-to-face communication frequencies rapidly decay with distance while telephone and email communication frequencies decay at slower rates.

When we analyzed the *propensity* to use each of the three media, we found that the use of face-to-face communication is substituted by telephone and email communication when distance increases. Furthermore, our empirical evidence shows (see Figure 6.5) that the relative use of telephone communication starts to decay after around 3000 kms, possibly because time-zone difference makes synchronous communication more difficult to accomplish.

Exploring this further, we found that team members located in countries that do not share the same first language show higher probability of using email communication than telephone communication. This supports the hypothesis that people with *language differences* prefer using

[12] For a formal introduction to the Design Structure Matrix (DSM) and how it is used to analyze the dependency structure of development projects refer to Steward (1981) and Eppinger, *et al.* (1994)

written, asynchronous communication media, such as email, rather than verbal, synchronous communication media, such as telephone. We recommend managers to identify whether there is significant language difference between team members involved in critical interactions in order to facilitate asynchronous-written communication.

In summary, relative location of interacting team members influences both communication frequency and media choice. Even if face-to-face communication can be substituted by other electronic-based communication such as email, instant messaging, or video-conferencing, managers should be aware that communication frequency tends to decrease with distance, independent of the media used to communicate. However, managers have other elements, such as team interdependence and organizational bonds, to mitigate the negative effects due to geographic dispersion of development organizations.

6.7. Limitations and Future Research

The fairly large size of our sample and the diverse nature of the projects examined offer encouragement as to the general nature of our findings. However, like most empirical research, there are significant limitations in our study. Our unit of analysis is the interacting pair. We do not attempt to describe how distance affects the propensity to communicate, only the frequency of communication and relative frequency of media use *given* that two people communicate. Also, our study is cross-sectional, not longitudinal. Thus, the standard caveats apply in drawing conclusions as to situations where one or more of the independent variables are adjusted due to managerial control.

The nature of information technology is changing at an incredible speed. At the time of the field study (1995), despite the fact that all three MNCs were themselves at the confluence of the merging technologies of computer and telephony, none of the development teams used, to any significant extent, emerging communication media such as video-conferencing, desktop conferencing or other "intra-net-based" technologies. Thus, our study is mainly limited to the three primary forms of communication used at the time: face-to-face, telephone and email. More research needs to be done to understand better the trade-off between media richness and data-transmission efficiency of the various communication media now widely available for development teams.

Furthermore, our study did not effectively examine the moderating effects of the content of communication. Clearly, some types of content are better suited to distant-communication than others. It would be useful to examine whether distance reduces communication frequency across media *and* content, and whether content has a significant influence on media choice.

Finally, we have not studied in detail the effect of barriers due to information technology differences. Our results emphasize the importance of minimizing such barriers between critically interdependent team members. Communication barriers due to information technology differences, such as corporate firewalls and incompatible information systems, have to be overcome to facilitate electronic information transfer between interdependent team members. An interesting stream of future research is to study the various effects imposed by these types of communication barriers.

7. *Contributions and Future Work*

"Jamás un hombre que no prefirió su Patria y la sirvió fiélmente
pasa a la historia sino con un nombre obscurecido"
Marzo 1824
Mariscal, Antonio José de Sucre

This thesis describes a method for analyzing the coupling between the architecture of a product and the development organization that designs it. Our approach involves three steps. First, we identify and document the design interfaces between the physical components that comprise the architecture of a product. Second, we identify the occurrence of technical interaction between design teams. We assume that design teams must interact with each other to address the technical design interfaces. Finally, we compare the predicted team interactions corresponding to each design interface with the actual team interactions.

By studying the coupling of product architecture and organizational structure of a large commercial aircraft engine we were able to predict 83% of the coordination-type interactions between the design teams that participated in the detailed design period of the development process. We focused our analysis on explaining the mismatch between design interfaces and team interactions. We have contributed to an understanding of what drives technical communication in product development organizations by formulating and testing several hypotheses to explain the cases when: 1) known design interfaces were not matched by team interactions, and 2) reported team interactions were not predicted by design interfaces.

7.1. About the Research Method

The research method presented in this thesis illustrates a novel approach to study the coupling of two key domains of strategic importance for established firms, the product architecture domain and the development organization domain. By documenting the product architecture in the design interface matrix we capture the current structure of a product. Then, by documenting the technical interactions between design teams in the team interaction matrix we capture the integration efforts of the development organization. Finally, by comparing these two matrices to generate the resultant matrix we create a valuable framework that provides the roadmap to analyze how the product architecture drives technical communication in product development organizations.

First, we confirmed the nominal hypothesis that team interactions are not statistically independent of design interfaces. Indeed, there is a strong association between the occurrence of team interactions and the existence of design interfaces. Then, we focused on understanding the mismatch between design interfaces and team interactions. We formulated several hypotheses to explain the unexpected cases when: 1) design interfaces where not matched by team interactions (the "X" cells of the resultant matrix), and 2) team interactions were not predicted by design interfaces (the "O" cells of the resultant matrix).

Our research approach is particularly easy to implement when there is a one-to-one mapping between the product architecture and the development organization. We developed an algebraic model (see Appendix C) that allows one to determine a predicted team interaction matrix for the cases when there is not a direct mapping between physical components and design teams.

By studying the coupling of the architecture of an aircraft engine and its development organization we have gained important insights about how product architecture drives technical communication in product development organizations. While we cannot claim the generality of these findings before completing similar studies in other types of products in different industries, we would expect to find similar results in other projects developing complex systems and where the development teams are organized according to the product architecture.

7.2. About the Analysis

We provided two different ways to analyze the results captured in the resultant matrix:

7.2.1. Categorical data analysis

First, we assumed that every cell in both the design interface matrix and the team interaction matrix are not only statistically independent between the matrices but also within the matrices. That assumption allowed us to use categorical data analysis techniques to test the hypotheses formulated to explain the mismatch between design interfaces and team interactions.

Under the categorical data analysis we completed three types of tests. First, we used chi-square tests of independence to test the hypothesized effects presented in chapter three. Second, we used chi-square tests of homogeneity to test the moderating effects due to systems modularity. Finally, we constructed three-dimensional contingency tables to test the moderating effects of organizational/system boundaries.

The advantage of using categorical data analysis is that since the unit of analysis is the cell, we can select sub-groups of cells according to the hypothesized effect we want to test. At the same time, the main limitation of the analysis is that its modeling unit is the cell and therefore it assumes statistically independent cells.

7.2.2. The log-linear model

Many researchers have shown empirically that social network data possess strong deviation from randomness. Hence, we needed to validate the assumption of independent cells within the matrices considered when completing the categorical data analysis. We validated the independent-cell assumption by developing a log-linear model based on Holland and Leinhardt's (1981) p_1 distribution. This model allows for the simultaneous estimation of parameters that measure the average tendency toward *reciprocation* (ρ), and the amount of differential *expansiveness* and differential *attraction* of each element of the network (α's and β's, respectively). This model provides a more formal statistical method to test the association between two networks that follow the p_1 distribution.

We extended the base log-linear model to include structural terms that allowed us to simultaneously test the hypothesized effects of organizational/system boundaries and the moderating effects of systems modularity. The results showed that these effects were statistically significant confirming the findings obtained from the categorical data analysis.

We found empirical evidence that, similarly to social networks, network of product's components exhibit a p_1 distribution. Having developed this log-linear model, based on social network analysis, provides a more robust method to formally test the level of association between two networks that exhibit a p_1 distribution.

7.3. Designing Modular versus Integrative Systems

In very complex products, we applied the definitions of modular and integral architecture (Ulrich and Eppinger, 1995) at the level of the many systems (and subsystems) which comprise the product. We defined modular systems as those exhibiting modular architecture characteristics while integrative systems are those with features of integral architectures. By analyzing the distribution of design interfaces across system boundaries we formally identified modular and integrative systems. Studying how the hypothesized effects are moderated by systems modularity allows us to understand better the difference between designing modular versus integrative systems.

Our analysis provided three important findings:

1. The distributed nature of the integrative systems forces these design teams to overcome organizational barriers in order to handle design interfaces with all the systems. That is, effects of organizational barriers are more severe among teams that design modular systems.
2. Design interfaces across modular systems are more difficult for design experts to recognize than interfaces with integrative systems.
3. Design teams handle some design interfaces according to their type. We found that spatial-type design interfaces are largely addressed in the design of modular systems while transfer-type design interfaces are more likely to be handled in the design of integrative systems.

7.4. Managerial Implications

This thesis outlines a method for analyzing the architecture of a product, once it is known, to determine potential technical communication linkages in the development organization. We also address several issues that need to be taken into account when predicting communication linkages. Managers may be able to better use understanding of product architecture to design organizational structures effectively, which facilitate coordination-type communications. This further suggests that managers may be able to improve product development performance by effectively selecting team members to deal with specific critical design interfaces and by outlining organizational boundaries to foster critical technical teams interactions.

Our results suggest that organizational boundaries foster communication within boundaries and prevent communication across boundaries. Furthermore, the effects of organizational boundaries are even more dominant than the effects of design interface criticality. These results should be good news for managers because they can take direct actions to place the organizational boundaries so that their effects truly improve product development performance. Hence, by studying the product architecture managers should identify the critical design interfaces to conceive organizations whose boundaries enclose such critical interfaces.

Our results also suggest that managers should pay particular attention in identifying modular and integrative systems because the effects of organizational boundaries are moderated by systems modularity. Since modular systems are not perfectly modular, their cross-boundary design interfaces are less likely to be handled by design teams. On the other hand, design teams that develop integrative systems are less vulnerable to organizational boundaries due to the physically distributed nature of the systems they design. Hence, managers should identify the critical cross-boundary design interfaces occurring between modular systems to facilitate the technical interaction between the corresponding design teams.

We found empirical evidence that suggests that system integration teams (i.e. design teams that are not responsible for the design of any specific component, but rather are in charge for system-level design issues) exchange technical information with almost every design team in the organization. This further indicates that the use of system integration teams can be an effective mechanism to handle critical cross-boundary design interfaces.

The effects of system boundaries were highlighted by the existence of team interactions that were not predicted by design interfaces. Such empirical evidence provided great benefits to the organization where our approach was implemented. In particular, the development organization for the next model of that aircraft engine assigned a design team that would handle those cross-boundary design interfaces that had not been recognized before by the design experts.

We found empirical evidence that supports the hypothesis that indirect exchange of technical information between design teams may take place across organizational boundaries. Indirect team interactions mitigate the negative effects of organizational boundaries because they handle design interfaces that are not addressed by direct team interactions. However, indirect team interactions are difficult to plan. Nonetheless, use of electronic-based integration systems (i.e.

centralized databases or product-data management systems) provides the tools to facilitate cross-boundary indirect team interactions. Hence, managers can facilitate effective indirect interaction by promoting the use of electronic-based integration systems.

From a product innovation viewpoint, the project studied was a mix of modular and incremental innovation. However, the lessons learned on this study may help development organizations to face architectural innovation. By documenting the architecture of the product in a design interface matrix for every generation of product family, novel architectures can be quickly identified. Furthermore, by documenting the interactions between the design teams (team interaction matrix) to compare them with the potential interactions provided by the design interface matrix provides a systematic way to evaluate how development organizations manage architectural knowledge, a critical issue for firms facing architectural innovation (Henderson and Clark, 1990).

What if the development organization is geographically distributed? We found empirical evidence that confirms previous research about how distance negatively influence communication. However, our results suggest that managers can mitigate distance effects by identifying critical team interdependence, using organizational bonds, and use of electronic-based communication media. Again, studying the product architecture would help managers to identify critical team interdependence and design organizational structures that foster critical non-collocated interactions.

7.5. Future Research Directions

This thesis opens a new stream of research in the interface of product architecture and development organization. This study is based on the assumption of a direct mapping of product architecture and development organization. What if this were not the case? Which types of barriers are more severe (organizational or system barriers)? Is an organizational design that mirrors the architecture of the product a good one? Extending this method to study various mappings of product architectures and development organizations may help us to answer these research questions.

To better understand the implications of this study on architectural innovation it is worth exploring the evolution over time of both design interface matrix and team interaction matrix for several generations in a product family. We expect the massive use of electronic-based communication media will improve the efficiency and effectiveness of the process of documenting team interactions over periods of time.

From an analytical perspective, future research might benefit from development of models that consider triadic effects. Those models might be better suited to test the effects of indirect team interactions and secondary design dependencies.

If we envision how product development will be done a few years from now, it is not difficult to imagine geographically distributed teams developing models that are seamlessly integrated with others' models through web-based tools (Wallace, *et al.* (2000)). How would the architecture of the product map the network of design models? Would teams interact more or less given that their models are integrated? What can we learn from studying the mismatch between models' dependency and team interactions, or between model's dependency and design interfaces?

These questions can be addressed by using the approach outlined in this thesis and taking advantage of having design models and design teams electronically connected. By counting electronic-based transactions we can assess the interaction intensity between design teams. Similarly, by counting the iterations between the various models that comprise a product we can estimate the level of dependency between them. Additionally, by studying the models' dependencies related to the architecture of the product we could obtain the design interfaces between the components of the product.

References

1. Adler, P. S. (1995), "Interdepartmental Interdependence and Coordination: The Case of the Design/Manufacturing Interface," *Organization Science,* 6, pp 147-167.

2. Alexander, C. (1964), *Notes on the Synthesis of Form,* Harvard University Press, Cambridge, MA.

3. Allen, T. J. (1964), "The use of information channels in research and development proposal preparation," *Working paper, MIT Sloan School of Management,* No. 97-64.

4. Allen, T. J. (1977), *Managing the Flow of Technology,* Cambridge: MIT Press.

5. Allen, T. J. (1986), "Organizational Structure, Information Technology and R&D Productivity," *IEEE Transactions on Engineering Management,* EM-33, 4, pp 212-217.

6. Allen, T. J. (1997), "Architecture and Communication among Product Development Engineers," *Sloan Working Paper # 3983.* Sloan School of Management. M.I.T.

7. Allen, T. J. and O. Hauptman (1987), "The Influence of Communication Technologies on Organization Structure: A Conceptual Structure for Future Research," *Communication Research,* 14(5): 575-587.

8. Allen, T. J. and O. Hauptman (1990), "The Substitution of Communication Technologies for Organizational Structure in Research and Development," In Fulk, J. and Steinfield, C.W. (Eds.), *Organizations and Communication Technology*: 275-294. Newbury Park CA: Sage.

9. Allen, T. J., D. M. Lee and M. L. Tushman (1980), "R&D Performance as a Function of Internal Communication, Project Management, and the Nature of the Work," *IEEE Transactions on Engineering Management,* Vol. EM-27, no. 1, pp 2-12, February.

10. Anderson, C. J., S. Wasserman (1995), "Log-multiplicative models for valued social relations," *Sociological Methods ε Res,* 24 96-127.

11. Barczak, G. and D. Wilemon (1991), "Communication Patterns of New Product Development Team Leaders," *IEEE Transactions on Engineering Management.* 38, No. 2.

12. Bishop, Y. M., Fienberg, S. E., and Holland, P. W. (1975), *Discrete Multivariate Analysis: Theory and Practice.* Cambridge, MA: The MIT Press.

13. Carrascosa, M., Eppinger, S. D. & Whitney, D. E. (1998), "Using Design Structure Matrix to Estimate Product Development Time". *Proceedings of DETC'98.* ASME Design Engineering Technical Conferences. DETC98/DAC-6013.

14. Chen, J. L. and D. S. Bolon (1993), "The Management of Multinational R&D: A Neglected Topic in International Business Research," *Journal of International Business Studies* 1: 1-18.

15. Clark, K. B. and T. Fujimoto (1991), *Product Development Performance: Strategy, Organization, and Management in the World Auto Industry*, Boston, MA: Harvard Business School.

16. Cochran, W. G. (1952), "The $\chi 2$ Goodness of Fit," *Annals of Mathematical Statistics* 23:315-345.

17. Cochran, W.G. (1954), "Some Methods for Strengthening the Common $\chi 2$ Tests," *Biometrics* 10:417-451.

18. Conrath, D. W. (1973), "Communications Environment and Its Relationship to Organizational Structure," *Management Science*, 20B 586-603.

19. Crawford, J. L. and G. A. Haaland (1972), "Predecisional Information-seeking and Subsequent Conformity in the Social Influence Process," *Journal of Personality and Social Psychology,* 23 112-119.

20. Daft, R. L. and R. H. Lengel (1984), "Information Richness: A New Approach to Managerial Information Processing and Organization Design," In Staw, B. and Cummings, L.L. (Eds.), *Research in Organizational Behavior* 6: 191-233. Greenwich, CT:JAI.

21. Daft, R. L. and R. H. Lengel (1986), "Organizational Information Requirements, Media Richness and Structural Design," *Management Science*, 32(5): 554-571.

22. Daniel, W. W., and Terrell, J. C. (1995), *Business Statistics for Management and Economics*. Houghton Mifflin Company.

23. De Meyer, A. (1991), "Tech talk: How Managers are Stimulating Global R&D Communication," *Sloan Management Review,* Spring: 49-58.

24. De Meyer, A. (1993), "Internationalizing R&D Improves a Firm's Technical Learning," *Research Technology Management,* July-August: 42-49.

25. De Meyer, A. and A. Mizushima (1989), "Global R&D Management." *R&D Management,* 19 135-146.

26. Eppinger, S. D., D. E. Whitney, R. P. Smith and D. A. Gebala (1994), "A Model-Based Method for Organizing Tasks in Product Development," *Research in Engineering Design*, vol 6, no. 1, pp 1-13.

27. Eppinger, S. D. (1997), "A Planning Method for Integration of Large-Scale Engineering Systems". *International Conference on Engineering Design (ICED 97 Tampere).* August.

28. Fienberg, S.E., S. S. Wasserman (1981), "Categorical data analysis of single sociometric relations," S. Leinhardt, ed. *Sociological Methodology* 1981. Jossey-Bass, San Francisco, CA. 156-192.

29. Fienberg, S.E., M. M. Meyer, S. S. Wasserman (1985), "Statistical analysis of multiple sociometric relations," *J. Amer. Stat. Assoc.* 80 51-67.

30. Galbraith, J. (1973), *Designing Complex Organizations.* Addison Wesley.

31. Gebala, D. A. and Eppinger, S. D. (1991), "Methods for Analyzing Design Procedures". *ASME Conference on Design Theory and Methodology.* Miami, FL, September 1991, DE-Vol 31, pp. 227-233.

32. Goshal, S. and C. A. Barlett (1990), "The Multinational Corporation as Interorganizational Network," *Academy Management Review*, Vol. 15, no. 4, pp 603-625, October.

33. Granstrand, O., L. Håkanson and S. Sjölander (1993), "Internationalization of R&D–a Survey of Some Recent Research," *Research Policy,* 22 413-430.

34. Griffin, A. and Hauser, J. R. (1992), "Patterns of Communication Among Marketing, Engineering and Manufacturing – A Comparison Between Two New Product Teams". *Management Science*. 38: (3) 360-373. March.

35. Griffin, A. (1997), "PDMA Research in NPD: Updating Trends and Benchmarking Best Practices," *Journal of Product Innovation Management,* 14 429-458.

36. Gulati, R. K. and S. D. Eppinger (1996), "The Coupling of Product Architecture and Organizational Structure Decisions," *Working Paper, Sloan School of Management,* No. 151-96. MIT.

37. Ha, A. and E. Porteus (1995), "Optimal Timing of Reviews in the Concurrent Design for Manufacturability," *Management Science* Vol. 41, No. 9. September.

38. Haberman, S. (1977), "Log-linear Models for Frequency Tables with Small Expected Counts", *Annals of Statistics*, 5:1148-1169.

39. Haberman, S. J. (1981), "Comment on 'An Exponential Family of Probability Densities for Directed Graphs,' by P. W. Holland and S. Leinhardt", *Journal of the American Statistical Association*, 76, 60-62.

40. Hakanson, L. and R.Nobel (1993), "Determinants of Foreign R&D in Swedish Multinationals," *Research Policy,* 22: 397-411.

41. Harary, F., Norman, R. Z. and Cartwright, D. (1965), *Structural Models: An Introduction to the Theory of Directed Digraphs*. New York: Wiley.

42. Harary, F. (1969), *Graph Theory,* Addison-Wesley, Reading, Massachusetts.

43. Hauptman, O. (1986), "Influence of Task Type on the Relationship Between Communication and Performance: The Case of Software Development," *R&D Management,* 16 127-139.

44. Henderson, R. and Clark, K. (1990), "Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms". *Administrative Science Quarterly*, vol. 35, no. 1, pp. 9-30.

45. Hightower, R. T. and L. Sayeed (1996), "Effects of Communication Mode and Prediscussion Information Distribution Characteristics on Information Exchange in Groups," *Information Systems Research,* 7 451-465.

46. Hiltz, S. R. and M. Turoff (1978), *The Network Nation*, Reading, MA: Addison-Wesley.

47. Holland, P. W., and Leinhardt, S. (1978), "An Omnibus Test for Social Structure Using Triads". *Sociological Methods and Research*, 7, 227-256.

48. Holland, P.W., S. Leinhardt (1981), "An exponential family of probability distributions for directed graphs," *J. Amer. Stat. Assoc.* 76 33-51.

49. Iacobucci, D., N. Hopkins (1994), "Detection of experimental effects in social network analysis," *Social Networks*. 16 1-41.

50. Jaffe, A., M. Trajtenberg and R. Henderson (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics,* 108 577-598.

51. Jones, J. W., C. Saunders and R. McLeod (1988), "Information Media and Source Patterns Across Management Levels: A Pilot Study," *Journal of Management Information Systems*, 5, 3 (Winter), 71-84.

52. Katz, R. and T. J. Allen (1982), "Investigating the not Invented Here (NIH) Syndrome: A Look at the Performance, Tenure and Communication Patterns of 50 R&D Projects Groups", *R&D Management*,12,1, pp 7-19.

53. Keller, R. T. (1986), "Predictors of the Performance of Project Groups in R&D Organizations", *Academy of Management Journal,* 29, 715-726.

54. Keller, R. T. and W. E.Holland (1983), "Communications and Innovations in Research and Development Organizations," *Academy of Management Journal,* 29, 715-726.

55. Krishnan, V., S. D. Eppinger, and D. E. Whitney (1997), "A Model-based Framework to Overlap Product Development Activities," *Management Science*, 43, 437-451.

56. Leonard, D. A., P. A.Brands, A. Edmondson and J. Fenwick (1997), "Virtual Teams: Using Communications Technology to Manage Geographically Dispersed Development Groups", *Sense and Respond: Capturing Value in the Network Era* edited by S.P. Bradley and R.L. Nolan. Cambridge: Harvard Business School Press.

57. Loch, C.H. and C. Terwiesch (1998), "Communication and Uncertainty in Concurrent Engineering," *Management Science*, 44, 8, 1032-1048. August.

58. Markus, M. L. (1994), "Electronic Mail as the Medium of Managerial Choice," *Organizational Science*, 5, 4, 502-527.

59. Marril, T. (1980), "Time to Retire the Telephone?" *Datamation*, (August), 185-186.

60. McCord, K. R. and S. D. Eppinger (1993), "Managing the Integration Problem in Concurrent Engineering," *Working Paper Sloan School of Management*, no. 3594,M.I.T.

61. McDonough III, E. F. and K. B. Kahn (1996), "Using 'Hard' and 'Soft' Technologies for Global New Product Development,"*R&D Management*, 26, 3 pp. 241-253.

62. McDonough III, E. F., K. B. Kahn and A. Griffin (1999), "Managing Communication in Global Product Development Teams," *IEEE Transactions on Engineering Management*, 46, 4, pp 375-386.

63. Morelli, M.D., S. D. Eppinger and R. K.Gulati (1995), "Predicting Technical Communication in Product Development Organizations," *IEEE Transactions on Engineering Management*, 42,3, pp 215-222.

64. Moreno, J.L. (1934), *Who Shall Survive?* Washington, D.C. Nervous and Mental Disease Co.

65. Nobel, R. and J. Birkinshaw (1998), "Innovation in Multinational Corporations: Control and Communication Patterns in International R&D Operations," *Strategic Journal of Management*, 19 479-496.

66. Pahl, G. and Beitz, W., (Wallace, K., ed.) (1991), *Engineering Design, A Systematic Approach*, Springer-Verlag, New York.

67. Pimmler, T. U. and Eppinger, S. D. (1994), "Integration Analysis of Product Decompositions". *Proceedings of the ASME Design Theory and Methodology Conference (DTM'94)*, DE-Vol 68, pp. 343-351.

68. Pinto, M. B., J. K. Pinto and J. E. Prescott (1993), "Antecedents and Consequences of Project Team Cross-functional Cooperation," *Management Science* 39 1281-1297.

69. Rice, J. A. (1994), *Mathematical Statistics and Data Analysis.* Belmont, CA : Duxbury Press.

70. Rowles, C. M. (1999), "System Integration Analysis of a Large Commercial Aircraft Engine". *Master's Thesis in Engineering and Management*, Massachusetts Institute of Technology.

71. Schmitz, J., and J. Fulk (1991), "Organizational Colleagues, Media Richness, and Electronic Mail: A Test of the Social Influence Model of Technology Use," *Communication Research*, 18, 4 (August), 487-523.

72. Siegel, J., V. Dubrovsky, S. Kiesler and T. W. McGuire (1984), *Group Processes in Computer-mediated Communication,* Pittsburgh, PA: Committee on Social Science Research in Computing Working Paper Series, Carnegie-Mellon University.

73. Simon, H.A. (1981), *The Science of the Artificial.* Second edition, MIT Press, Cambridge, MA.

74. Smith, G. F. and Browne, G. J. (1993), "Conceptual Foundations of Design Problem Solving". *IEEE Transactions on System, Man, and Cybernetics.* Vol. 23, No. 5, September/October.

75. Smith, R. P. and S. D. Eppinger (1997), "Identifying Controlling Features of Engineering Design Iteration" *Management Science*,43, 3, pp 276-293.

76. Sosa, M. E., Eppinger, S. D. and Pich, M. (2000), "Factors that Influence Technical Communication in Distributed Product Development: An Empirical Study in the Telecommunications Industry". *Sloan Working Paper*, MIT School of Management.

77. Sproull, L and S. Kiesler (1991), *Connections: New ways of working in the networked organization,* Cambridge: MIT Press.

78. Steward, D. (1981), "The Design Structure Matrix: A Method for Managing the Design of Complex Systems," *IEEE Transactions on Engineering Management,* EM-28, 3, 71-74.

79. Strang, G. (1980), *Linear Algebra and Its Applications,* New York: Academic Press.

80. Suh, N. P. (1990), *The Principles of Design,* Oxford University Press, New York.

81. Trevino, L. K., R. Lengel and R. L. Daft (1987), "Media Symbolism, Media Richness, and Media Choice in Organizations: A Symbolic Interactionist Perspective," *Communication Research,* 14 553-574.

82. Tushman, M. L. (1978), "Technical Communication in R&D Laboratories: The Impact of Project Work Characteristics," *Acad. Management Journal,* 21, 624-645.

83. Ulrich, K. T. (1995), "The Role of Product Architecture in the Manufacturing Firm". *Research Policy,* vol. 24, pp. 583-607.

84. Ulrich, K. T. and S. D. Eppinger (1995), *Product Design and Development.* McGraw Hill, Inc.

85. Van den Bulte, C. and R. K. Moenaert (1998), "The Effects of R&D Team Co-location on Communication Patterns among R&D, Marketing, and Manufacturing," *Management Science,* 44: S1-S19.

86. von Hippel, E. (1990), "Task Partitioning: An Innovative Process Variable". *Research Policy,* vol. 9, pp. 407-418.

87. Wageman, R. (1995), "Interdependence and Group Effectiveness," *Administrative Science Quarterly,* 40 145-180.

88. Wallace, D. R., S. Abrahamson, N. Senin, and P. Sferro (2000), "Integrated Design in a Service Marketplace", *Computer-aided Design,* vol. 32, no 2, pp. 97-107.

89. Wasserman, S. (1987), "Conformity of Two Sociometric Relations". *Psychometrika.* 52, 3-18.

90. Wasserman, S. and D.Iacobucci (1988), "Sequential Social Network Data," *Psychometrika*, 53 261-282.

91. Wasserman, S., K, Faust. (1994), *Social Network Analysis*. Cambridge University Press, Cambridge.

92. Wasserman, S., P. Pattison (1996), "Logit models and logistic regression for social networks: I. an introduction to markov graphs and p*," *Psychometrika*, 61, 401-425.

93. Wells, B. W. (1965), "The Psycho-social Influence of Building Environment: Sociometric Findings in Large and Small Office Spaces," *Building Science* 1 153-165.

94. Wheelwright, S. C. and K. B. Clark (1992), *Revolutionizing Product Development. Quantum Leaps in Speed, Efficiency, and Quality*, New York: Free Press.

95. Yates, J. and W. J. Orlikowski (1992), "Genres of Organizational Communication: An Approach to Studying Communication and Media," *Academy of Management Review*, 17, 299-326.

96. Zahn, G. L. (1991), "Face-to-face Communication in an Office Setting," *Communication Research* 18 737-754.

A. Chi-square Tests Results

Table A-1. Chi-square test of independence. Effect of design interface strength.

	Total	Expected number (fraction) of design interfaces matched by team interactions	Expected number (fraction) of design interfaces not matched by team interactions	Actual number (fraction) of design interfaces matched by team interactions	Actual number (fraction) of design interfaces not matched by team interactions	χ^2 of design interfaces matched by team interactions	χ^2 of design interfaces not matched by team interactions
Weak design interface (strength ≤4)	319	191 176 (61 34%)	127.824 (38 66%)	169 (52 98%)	150 (47.02%)	3 633	5 763
Strong design interface (strength >4)	250	149.824 (61.34%)	100 176 (38 66%)	180 (72.00%)	70 (28.00%)	4 635	7 354
Total	569	349.000	220.000	349	220	**8.268**	**13.116**

H_o: Weak design interfaces are as likely to be matched by team interactions as strong design interfaces.

χ^2 **= 21.385** Critical $\chi^2_{(0.99,1)} = 6.635$ Since χ^2 > Critical $\chi^2_{(0.99,1)}$, we reject H_o.

Table A2. Chi-square test of homogeneity. Effects of design interface strength

	Total	Expected number (fraction) of design interfaces matched by team interactions	Expected number (fraction) of design interfaces not matched by team interactions	Actual number (fraction) of design interfaces matched by team interactions	Actual number (fraction) of design interfaces not matched by team interactions	χ^2 of design interfaces matched by team interactions	χ^2 of design interfaces not matched by team interactions
Weak design interfaces (Modular systems)	140	74.169 (52.98%)	65.831 (47.02%)	70 (50.00%)	70 (50 00%)	0.234	0.264
Weak design interfaces (Integrative systems)	179	94.831 (52.98%)	84.169 (47.02%)	99 (55.31%)	80 (44.69%)	0 183	0 207
Total	319	169.000	150.000	169	150	**0.418**	**0.471**
Strong design interfaces (Modular systems)	107	77.040 (72 00%)	29.960 (28.00%)	80 (74 77%)	27 (25.23%)	0.478	0.439
Strong design interfaces (Integrative systems)	143	102 960 (72.00%)	40.040 (28.00%)	100 (69.93%)	43 (30.07%)	0.871	0 799
Total	250	180.000	70.000	180	70	**0.199**	**0.511**

H_o: The effect of design interface strength is the same on modular systems as on integrative systems.

$\chi^2_{overall}$ **= 1.598** Critical $\chi^2_{(0.99,3)} = 11.345$ Since χ^2 < Critical $\chi^2_{(0.99,3)}$, we do not reject H_o

χ^2_{weak} **= 0.888** Critical $\chi^2_{(0.99,1)} = 6.635$ Since χ^2 < Critical $\chi^2_{(0.99,1)}$, we do not reject H_o

χ^2_{strong} **= 0.710** Critical $\chi^2_{(0.99,1)} = 6.635$ Since χ^2 < Critical $\chi^2_{(0.99,1)}$, we do not reject H_o.

Table A3. Chi-square test of independence of organizational boundaries and design interface strength

	Total	Expected number (fraction) of design interfaces within organizational boundaries	Expected number (fraction) of design interfaces across organizational boundaries	Actual number (fraction) of design interfaces within organizational boundaries	Actual number (fraction) of design interfaces across organizational boundaries	χ^2 of design interfaces within organizational boundaries	χ^2 of design interfaces across organizational boundaries
Weak design interface	319	129.506 (40.60%)	189.494 (59.40%)	95 (30.09%)	223 (69.91%)	8.669	5.925
Strong design interface	250	101.494 (40.60%)	148.506 (59.40%)	135 (54.00%)	115 (46.00%)	11.061	7.560
Total	569	231.000	338.000	231	338	**19.730**	**13.484**

H_o: The strength of the design interface is independent of whether or not the design interface is within or across organizational boundaries.

χ^2 = **33.214** Critical $\chi^2_{(0.99,1)} = 6.635$ Since χ^2 > Critical $\chi^2_{(0.99,1)}$, we reject H_o.

Table A4. Chi-square test of independence. Effect of design interface type.

	Total	Expected number (fraction) of design interfaces matched by team interactions	Expected number (fraction) of design interfaces not matched by team interactions	Actual number (fraction) of design interfaces matched by team interactions	Actual number (fraction) of design interfaces not matched by team interactions	χ^2 of design interfaces matched by team interactions	χ^2 of design interfaces not matched by team interactions
Spatial-type	41	15.459 (37.70%)	25.541 (62.30%)	20 (48.78%)	21 (51.22%)	1.334	0.807
Transfer-type	81	30.541 (37.70%)	50.459 (62.30%)	26 (32.10%)	55 (67.90%)	0.675	0.409
Total	122	46.000	76.000	46	76	**2.009**	**1.216**

H_o: Spatial-type design interfaces are as likely to be matched by team interactions as transfer-type design interfaces.

χ^2 = **3.225** Critical $\chi^2_{(0.95,1)} = 3.841$ Since χ^2 < Critical $\chi^2_{(0.95,1)}$, we do not reject H_o.

Table A5. Chi-square test of homogeneity. Effect of design interface type.

	Total	Expected number (fraction) of design interfaces matched by team interactions	Expected number (fraction) of design interfaces not matched by team interactions	Actual number (fraction) of design interfaces matched by team interactions	Actual number (fraction) of design interfaces not matched by team interactions	χ^2 of design interfaces matched by team interactions	χ^2 of design interfaces not matched by team interactions
Spatial-type (Modular systems)	24	11.707 (48.78%)	12.293 (51.22%)	15 (62 50%)	9 (37.50%)	0.926	0.882
Spatial-type (Integrative systems)	17	8.293 (48.78%)	8.707 (51.22%)	5 (29.41%)	12 (70.59%)	1.307	1.245
Total	41	20.0	21.0	20	21	2.233	2.127
Transfer-type (Modular systems)	41	13.160 (32.10%)	27.840 (67.90%)	8 (19.51%)	33 (80.49%)	2.024	0.957
Transfer-type (Integrative systems)	40	12.840 (32 10%)	27.160 (67.90%)	18 (45.00%)	22 (55.00%)	2 074	0.980
Total	81	26.0	55 0	26	55	4.098	1.937

H_o: The effect of design interface type is the same on modular systems than on integrative systems.

$\chi^2_{overall}$ **= 10.395.** Critical $\chi^2_{(0.95,3)} = 7.815$. Since χ^2 > Critical $\chi^2_{(0.95,3)}$, we reject H_o.

$\chi^2_{spatial\text{-}type}$ **= 4.360** Critical $\chi^2_{(0.95,1)} = 3.841$. Since χ^2 > Critical $\chi^2_{(0.95,3)}$, we reject H_o.

$\chi^2_{transfer\text{-}type}$ **= 6.035** Critical $\chi^2_{(0.95,1)} = 3.841$. Since χ^2 > Critical $\chi^2_{(0.95,3)}$, we reject H_o.

Table A6. Chi-square test of independence.
Effect of indirect team interactions (through other design teams)

	Total	Expected number (fraction) of indirect team interactions covering design interfaces	Expected number (fraction) of indirect team interactions not covering design interfaces	Actual number (fraction) of indirect team interactions covering design interfaces	Actual number (fraction) of indirect team interactions not covering design interfaces	χ^2 of indirect team interactions covering design interfaces	χ^2 of indirect team interactions not covering design interfaces
No potential indirect team interaction	1401	126.371 (9.02%)	1274.629 (90.98%)	70 (5.00%)	1331 (95.00%)	25.146	2.493
One potential indirect team interaction	579	52.226 (9.02%)	526.774 (90.98%)	59 (10.19%)	520 (89.81%)	0.879	0.087
More than one potential indirect team interaction	459	41.402 (9.02%)	417.598 (90.98%)	91 (19.83%)	368 (80.17%)	59.416	5.891
Total	2439	220.000	2219.000	220	2219	**85.440**	**8.471**

H_o: The likelihood that two design teams, who do not interact directly, share a design interface is independent of the number of potential indirect interactions through other design teams.

χ^2 **= 93.911** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 > Critical $\chi^2_{(0.99,2)}$, we reject H_o.

Table A7. Chi-square test of homogeneity.
Effects of indirect team interactions (through other design teams).

	Total	Expected number (fraction) of indirect team interactions covering design interfaces	Expected number (fraction) of indirect team interactions not covering design interfaces	Actual number (fraction) of indirect team interactions covering design interfaces	Actual number (fraction) of indirect team interactions not covering design interfaces	χ^2 of indirect team interactions covering design interfaces	χ^2 of indirect team interactions not covering design interfaces
No indirect interaction (Modular systems)	747	37.323 (5.00%)	709.677 (95.00%)	29 (3.88%)	718 (96.12%)	1.856	0.098
One indirect interaction (Modular systems)	258	26.290 (10.19%)	231.710 (89.81%)	31 (12.02%)	227 (87.98%)	0.844	0.096
More than one indirect interaction (Modular systems)	138	27.359 (19.83%)	110.641 (80.17%)	37 (26.81%)	101 (73.19%)	3.397	0.840
No indirect interaction (Integrative systems)	654	32.677 (5.00%)	621.323 (95.00%)	41 (6.27%)	613 (93.73%)	2.120	0.112
One indirect interaction (Integrative systems)	321	32.710 (10.19%)	288.290 (89.81%)	28 (8.72%)	293 (91.28%)	0.678	0.077
More than one indirect interaction (Integrative systems)	321	63.641 (19.83%)	257.359 (80.17%)	54 (16.82%)	267 (83.18%)	1.460	0.361
Total	2439	220.000	2219.000	220	2219	**10.356**	**1.583**

H_o: The effect of potential indirect interactions is the same for modular and integrative systems.

χ^2_{overall} = **11.939** Critical $\chi^2_{(0.99,5)} = 15.086$ Since χ^2 < Critical $\chi^2_{(0.99,5)}$, we do not reject H_o

Table A8. Chi-square test of independence.
Indirect team interactions (through other design teams) vs. organizational boundaries

	Total	Expected number (fraction) of indirect team interactions within organizational boundaries	Expected number (fraction) of indirect team interactions across organizational boundaries	Actual number (fraction) of indirect team interactions within organizational boundaries	Actual number (fraction) of indirect team interactions across organizational boundaries	χ^2 of indirect team interactions within organizational boundaries	χ^2 of indirect team interactions across organizational boundaries
No indirect team interactions	1401	68.930 (4.92%	1332.070 (95.08%)	28 (2.00%)	1373 (98.00%)	24.304	1.258
One indirect team interactions	579	28.487 (4.92%)	550.513 (95.08%)	31 (5.35%)	548 (94.65%)	0.222	0.011
More than one indirect team interactions	459	22.583 (4.92%)	436.417 (95.08%)	61 (13.29%)	398 (86.71%)	65.353	3.382
Total	2439	120.000	2319.000	120	2319	**89.878**	**4.651**

H_o: The number of potential indirect team interactions between two teams is independent of whether they are within or across organizational boundaries.

χ^2 = **94.529** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 > Critical $\chi^2_{(0.99,2)}$, we reject H_o.

Table A9. Chi-square test of homogeneity. Effects of organizational boundaries controlling for indirect team interactions (through other design teams).

	Total	Expected number (fraction) of indirect team interactions covering design interfaces	Expected number (fraction) of indirect team interactions not covering design interfaces	Actual number (fraction) of indirect team interactions covering design interfaces	Actual number (fraction) of indirect team interactions not covering design interfaces	χ^2 of indirect team interactions covering design interfaces	χ^2 of indirect team interactions not covering design interfaces
No indirect interaction (Within organizational boundaries)	28	10.267 (36.67%)	17.733 (63.33%)	15 (53.57%)	13 (46.43%)	2.182	1.263
One indirect interaction (Within organizational boundaries)	31	11.367 (36.67%)	19.633 (63.33%)	7 (22.58%)	24 (77.42%)	1.678	0.971
More than one indirect interaction (Within organizational boundaries)	61	22.367 (36.67%)	38.633 (63.33%)	22 (36.07%)	39 (63.93%)	0.006	0.003
Total	120	44.000	76.000	44	76	**3.866**	**2.238**
No indirect interaction (Across organizational boundaries)	1373	104.204 (7.59%)	1268.796 (92.41%)	55 (4.01%)	1318 (95.99%)	23.233	1.908
One indirect interaction (Across organizational boundaries)	548	41.590 (7.59%)	506.410 (92.41%)	52 (9.49%)	496 (90.51%)	2.605	0.214
More than one indirect interaction (Across organizational boundaries)	398	30.206 (7.59%)	367.794 (92.41%)	69 (17.34%)	329 (82.66%)	49.823	4.092
Total	2319	176.000	2143.000	176	2143	**75.662**	**6.214**

H_o: The effect of potential indirect interactions is the same within or across organizational boundaries.

χ^2_{overall} = **87.980** Critical $\chi^2_{(0.99,5)} = 15.086$ Since χ^2 > Critical $\chi^2_{(0.99,5)}$, we reject H_o.

χ^2_{within} = **6.104** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 < Critical $\chi^2_{(0.99,2)}$, we do not reject H_o.

χ^2_{across} = **81.876** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 > Critical $\chi^2_{(0.99,2)}$, we reject H_o.

Table A10. Chi-square test of homogeneity. Joint effects of organizational boundaries and indirect team interactions (through other design teams).

	Total	Expected number (fraction) of indirect team interactions covering design interfaces	Expected number (fraction) of indirect team interactions not covering design interfaces	Actual number (fraction) of indirect team interactions covering design interfaces	Actual number (fraction) of indirect team interactions not covering design interfaces	χ^2 of indirect team interactions covering design interfaces	χ^2 of indirect team interactions not covering design interfaces
No indirect interaction (Within organizational boundaries)	28	2.526 (9.02%)	25.474 (90.98%)	15 (53.57%)	13 (46.43%)	61.612	6.108
One indirect interaction (Within organizational boundaries)	31	2.796 (9.02%)	28.204 (90.98%)	7 (22.58%)	24 (77.42%)	6 320	0 627
More than one indirect interaction (Within organizational boundaries)	61	5.502 (9 02%)	55.498 (90.98%)	22 (36.07%)	39 (63.93%)	49 466	4.904
Total	120			44	76	**117.399**	**11.639**
No indirect interaction (Across organizational boundaries)	1373	123.846 (9.02%)	1249.154 (90.98%)	55 (4.01%)	1318 (95.99%)	38.271	3 794
One indirect interaction (Across organizational boundaries)	548	49.430 (9.02%)	498.570 (90.98%)	52 (9.49%)	496 (90.51%)	0.134	0 013
More than one indirect interaction (Across organizational boundaries)	398	35.900 (9.02%)	362.100 (90.98%)	69 (17.34%)	329 (82.66%)	30.518	3 026
Total	2319			176	2143	**68.923**	**6.833**

H_o: The joint effect of organizational boundaries and potential indirect interactions is independent of whether design interfaces are matched by team interactions or not.

$\chi^2_{overall}$ = **204.795** Critical $\chi^2_{(0.99,5)} = 15.086$ Since $\chi^2 >$ Critical $\chi^2_{(0.99,5)}$, we reject H_o.

Table A11. Chi-square test of independence. Effect of indirect team interactions (through system integration teams)

	Total	Expected number (fraction) of indirect team interactions covering design interfaces	Expected number (fraction) of indirect team interactions not covering design interfaces	Actual number (fraction) of indirect team interactions covering design interfaces	Actual number (fraction) of indirect team interactions not covering design interfaces	χ^2 of indirect team interactions covering design interfaces	χ^2 of indirect team interactions not covering design interfaces
No indirect interactions with system integrators	1098	99.041 (9.02%)	998.959 (90.98%)	91 (8.29%)	1007 (91.71%)	0.653	0.065
One indirect interaction with system integrators	1044	94.170 (9.02%)	949.830 (90.98%)	102 (9.77%)	942 (90.23%)	0.651	0.065
More than one indirect interaction with system integrators	397	26.790 (9.02%)	270.210 (90.98%)	27 (9.09%)	270 (90.91%)	0.002	0.000
Total	2439	220.000	2219.000	220	2219	**1.306**	**0.129**

H_o: The likelihood that two design teams, who do not interact directly, share a design interface is independent of the number of indirect interactions through system integration teams.

χ^2 **= 1.435** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 < Critical $\chi^2_{(0.99,2)}$, we do not reject H_o.

Table A12. Chi-square test of independence. Effect of system boundaries

	Total	Expected number (fraction) of predicted team interactions	Expected number (fraction) of unpredicted team interactions	Actual number (fraction) of predicted team interactions	Actual number (fraction) of unpredicted team interactions	χ^2 of predicted team interactions	χ^2 of unpredicted team interactions
Team interactions within system boundaries	208	171.612 (82.51%)	36.388 (17.49%)	187 (89.90%)	21 (10.10%)	1.380	6.507
Team interactions across system boundaries	215	177.388 (82.51%)	37.612 (17.49%)	162 (75.35%)	53 (24.65%)	1.335	6.295
Total	423	349.000	74.000	349	74	**2.715**	**12.803**

H_o: Team interactions within system boundaries are as likely to be predicted design interfaces as team interactions across system boundaries.

χ^2 **= 15.517** Critical $\chi^2_{(0.99,1)} = 6.635$ Since χ^2 > Critical $\chi^2_{(0.99,1)}$, we reject H_o.

Table A-13. Chi-square test of homogeneity. Effects of system boundaries.

	Total	Expected cases of predicted team interactions	Expected cases of unpredicted team interactions	Actual cases of predicted team interactions	Actual cases of unpredicted team interactions	χ2 of predicted team interactions	χ2 of unpredicted team interactions
Design interfaces within system boundaries (Modular systems)	124	111.5 (89.9%)	12.5 (10.1%)	110 (88.7%)	14 (11.3%)	0.020	0.175
Design interfaces within system boundaries (Integrative systems)	84	75.5 (89.9%)	8.5 (10.1%)	77 (91.7%)	7 (8.3%)	0.029	0.259
Total	208	187.0	21.0	187	21	**0.049**	**0.434**
Design interfaces across system boundaries (Modular systems)	65	49.0 (75.3%)	16.0 (24.7%)	40 (61.5%)	25 (38.5%)	1.645	5.029
Design interfaces across system boundaries (Integrative systems)	150	113.0 (75.3%)	37.0 (24.7%)	122 (81.3%)	28 (18.7%)	0.713	2.179
Total	215	162.0	53.0	162	53	**2.358**	**7.208**

H_o: The effect of system boundaries is the same for modular and integrative systems.

$\chi^2_{\text{within-boundaries}} = 0.483$ $\chi^2_{\text{across-boundaries}} = 9.566$ $\chi^2_{\text{critical (0.95,1)}} = 3.841$

Table A14. Chi-square test of independence. Effect of secondary design dependencies.

	Total	Expected number (fraction) of secondary design interfaces matched by team interactions	Expected number (fraction) of secondary design interfaces not matched by team interactions	Actual number (fraction) of secondary design interfaces matched by team interactions	Actual number (fraction) of secondary design interfaces not matched by team interactions	χ^2 of secondary design interfaces matched by team interactions	χ^2 of secondary design interfaces not matched by team interactions
No secondary design interfaces	1005	32.433 (3.23%)	972.567 (96.77%)	5 (0.50%)	1000 (99.50%)	23.204	0.774
Three or fewer secondary design interfaces	643	20.751 (3.23%)	622.249 (96.77%)	25 (3.89%)	618 (96.11%)	0.870	0.029
More than three secondary design interfaces	645	20.816 (3.23%)	624.184 (96.77%)	44 (6.82%)	601 (93.18%)	25.823	0.861
Total	2293	74.000	2219.000	74	2219	**49.897**	**1.664**

H_o: The likelihood that two design teams, who do not share direct design interfaces, interact is independent of the number of secondary design dependencies.

χ^2 **= 51.561** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 > Critical $\chi^2_{(0.99,2)}$, we reject H_o.

Table A15. Chi-square test of homogeneity. Effects of system boundaries controlling for secondary design dependency.

	Total	Expected number (fraction) of secondary design interfaces matched by team interactions	Expected number (fraction) of secondary design interfaces not matched by team interactions	Actual number (fraction) of secondary design interfaces matched by team interactions	Actual number (fraction) of secondary design interfaces not matched by team interactions	χ^2 of secondary design interfaces matched by team interactions	χ^2 of secondary design interfaces not matched by team interactions
No secondary dependency (Modular systems)	594	2.955 (0.50%)	591.045 (99.50%)	4 (0.67%)	590 (99.33%)	0.369	0.002
Three or fewer secondary dependencies (Modular systems)	259	10.070 (3.89%)	248.930 (96.11%)	17 (6.56%)	242 (93.44%)	4.769	0.193
More than three secondary dependencies (Modular systems)	232	15.826 (6.82%)	216.174 (93.18%)	18 (7.76%)	214 (92.24%)	0.299	0.022
No secondary dependency (Integrative systems)	411	2.045 (0.50%)	408.955 (99.50%)	1 (0.24%)	410 (99.76%)	0.534	0.003
Three or fewer secondary dependencies (Integrative systems)	384	14.930 (3.89%)	369.070 (96.11%)	8 (2.08%)	376 (97.92%)	3.217	0.130
More than three secondary dependencies (Integrative systems)	413	28.174 (6.82%)	384.826 (93.18%)	26 (6.30%)	387 (93.70%)	0.168	0.012
Total	2293	74.000	2219.000	74	2219	**9.355**	**0.362**

H_o: The effect due to secondary design dependencies is the same for modular and integrative systems.

$\chi^2_{overall}$ = **9.717** Critical $\chi^2_{(0.99,5)} = 15.086$ Since χ^2 < Critical $\chi^2_{(0.99,5)}$, we do not reject H_o.

Table A16. Chi-square test of independence. Secondary design dependencies vs. System boundaries.

	Total	Expected number (fraction) of secondary design interfaces within system boundaries	Expected number (fraction) of secondary design interfaces across system boundaries	Actual number (fraction) of secondary design interfaces within system boundaries	Actual number (fraction) of secondary design interfaces across system boundaries	χ^2 of secondary design interfaces within system boundaries	χ^2 of secondary design interfaces across system boundaries
No secondary design interfaces	1005	42.514 (4.23%)	962.486 (95.77%)	2 (0.20%)	1003 (99.80%)	38.608	1.705
Three or fewer secondary design interfaces	643	27.201 (4 23%)	615.799 (95.77%)	29 (4.51%)	614 (95.49%)	0.119	0.005
More than three secondary design interfaces	645	27.285 (4.23%)	617.715 (95.77%)	66 (10.23%)	579 (89.77%)	54 932	2.426
Total	2293	97.000	2196.000	97	2196	**93.659**	**4.137**

H_o: The number of secondary design dependencies between two components is independent of whether they belong to the same system or not.

χ^2 = **97.796** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 > Critical $\chi^2_{(0.99,2)}$, we reject H_o.

Table A17. Chi-square test. Effects of secondary design dependency controlling for system boundaries.

	Total	Expected number (fraction) of secondary design interfaces matched by team interactions	Expected number (fraction) of secondary design interfaces not matched by team interactions	Actual number (fraction) of secondary design interfaces matched by team interactions	Actual number (fraction) of secondary design interfaces not matched by team interactions	χ^2 of secondary design interfaces matched by team interactions	χ^2 of secondary design interfaces not matched by team interactions
No secondary dependency (Within boundaries)	2	0.433 (21.65%)	1.567 (78.35%)	0 (0.00%)	2 (100.00%)	0.433	0.120
Three or fewer secondary dependencies (Within boundaries)	29	6.278 (21.65%)	22.722 (78.35%)	8 (27.59%)	21 (72.41%)	0.472	0.130
More than three secondary dependencies (Within boundaries)	66	14.289 (21.65%)	51.711 (78.35%)	13 (19.70%)	53 (80.30%)	0.116	0.032
Total	97	21.000	76.000	21	76	**1.021**	**0.282**
No secondary dependency (Across boundaries)	1003	24.207 (2.41%)	978.793 (97.59%)	5 (0.50%)	998 (99.50%)	15.240	0.377
Three or fewer secondary dependencies (Across boundaries)	614	14.819 (2.41%)	599.181 (97.59%)	17 (2.77%)	597 (97.23%)	0.321	0.008
More than three secondary dependencies (Across boundaries)	579	13.974 (2.41%)	565.026 (97.59%)	31 (5.35%)	548 (94.65%)	20.744	0.513
Total	2196	53.000	2143.000	53	2143	**36.305**	**0.898**

Ho: The likelihood that two design teams, who do not share direct design interfaces, interact is independent of the number of secondary design dependencies (for both within and across system boundaries)

$\chi^2_{overall}$ = **38.507** Critical $\chi^2_{(0.99,5)} = 15.086$ Since χ^2 > Critical $\chi^2_{(0.99,5)}$, we reject H_o

χ^2_{within} = **1.304** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 < Critical $\chi^2_{(0.99,2)}$, we do not reject H_o

χ^2_{across} = **37.203** Critical $\chi^2_{(0.99,2)} = 9.210$ Since χ^2 > Critical $\chi^2_{(0.99,2)}$, we reject H_o.

Table A18. Chi-square test. Joint effects of system boundaries and secondary design dependency.

	Total	Expected number (fraction) of secondary design interfaces matched by team interactions	Expected number (fraction) of secondary design interfaces not matched by team interactions	Actual number (fraction) of secondary design interfaces matched by team interactions	Actual number (fraction) of secondary design interfaces not matched by team interactions	χ^2 of secondary design interfaces matched by team interactions	χ^2 of secondary design interfaces not matched by team interactions
No secondary dependency (Within boundaries)	2	0.065 (3.23%)	1.935 (96.77%)	0 (0.00%)	2 (100.00%)	0.065	0.002
Three or fewer secondary dependencies (Within boundaries)	29	0.936 (3.23%)	28.064 (96.77%)	8 (27.59%)	21 (72.41%)	53.320	1.778
More than three secondary dependencies (Within boundaries)	66	2.130 (3.23%)	63.870 (96.77%)	13 (19.70%)	53 (80.30%)	55.474	1.850
Total	97			21	76	**108.859**	**3.630**
No secondary dependency (Across boundaries)	1003	32.369 (3.23%)	970.631 (96.77%)	5 (0.50%)	998 (99.50%)	23.141	0.772
Three or fewer secondary dependencies (Across boundaries)	614	19.815 (3.23%)	594.185 (96.77%)	17 (2.77%)	597 (97.23%)	0.400	0.013
More than three secondary dependencies (Across boundaries)	579	18.686 (3.23%)	560.314 (96 77%)	31 (5.35%)	548 (94.65%)	8.116	0 271
Total	2196			53	2143	**31.657**	**1.056**

H_o: The joint effect of secondary design dependencies and system boundaries is independent of whether team interactions are reported or not.

$\chi^2_{overall}$ **= 145.201** Critical $\chi^2_{(0.99,5)} = 15.086$ Since χ^2 > Critical $\chi^2_{(0.99,5)}$, we reject H_o

B. The Effects due to Interface Redesign

This appendix includes the description of the statistical analyses, and additional data collected to test the effects of interface redesign (hypothesis H3.4).

B.1. Regression Analysis

Linear regression between the percentage of redesign of component i and the fraction of design interfaces of component i matched by team interactions. That is, we test whether the correlation between the estimated percentage of redesign of component i and the fraction of design interfaces of component i matched by team interactions (i.e. ratio of number of "#" cells in row i of the resultant matrix over the number of "X" cells in row i of the binary design interface matrix) was positively statistically significant. The test resulted in a not statistically significant correlation coefficient equal to 0.08 (p-value = 0.26).

B.2. Change and Impact of the interface

Let us consider the design interface in which component i depends on component j. Previous researchers (Krishnan and Eppinger (1997), Loch and Terwiesch (1998), Carrascosa *et al.* (1999)) have used the concepts of evolution and sensitivity to categorize information transfer between development tasks. We extend those concepts to define the redesign of an interface. We define *change of* component j, and *impact on* component i as follows:

- **Change** of component j (p_j) is the probability that new design information that affects its interface with component i would be generated.
- **Impact** on component i (r_i) is the probability that changes in the interface with component j would generate design changes in component i.

B.3. Additional Data Collection

Based on the concepts of change and impact of the interface, we collect additional information to capture the redesign between the design interfaces of the high-pressure turbine (HPT) and low-pressure turbine (LPT).

We document the interface redesign data in the change interface matrix (Figure B.1) and in the impact interface matrix (Figure B.2). A mark in row *i* of the change interface matrix indicates that redesign of component *i* generated changes to the interface with component *j*. Similarly, a mark in row *i* of the impact interface matrix indicates that changes in the interface with component *j* had an impact on the design of component *i*.

		28	29	30	31	32	33	34	35	36	37	38
HPT Blades	28	*		C	C					C	C	
HPT CV	29	C	*	C		C						
HPT 2V	30	C		*	C	C						
HPT Rotor	31	C			*	C						C
HPT Case/OAS	32			C		*		C				
LP Shaft	33				C		*					
LPT Case	34							*		C		
TEC	35								*			
LPT Vanes	36							C		*	C	
LPT Blades	37										*	
LPT OAS / TDucts / Insulation	38											*

Figure B-1. Change Interface Matrix (binary)

		28	29	30	31	32	33	34	35	36	37	38
HPT Blades	28	*	I	I	I	I						
HPT CV	29		*									
HPT 2V	30	I		*	I							
HPT Rotor	31	I		I	*							
HPT Case/OAS	32			I	I	*						
LP Shaft	33						*					
LPT Case	34					I		*		I		
TEC	35								*			
LPT Vanes	36							I		*		
LPT Blades	37									I	*	
LPT OAS / TDucts / Insulation	38											*

Figure B-2. Impact Interface Matrix (binary)

B.4. Chi-square Tests Results

We analyze the 57 design interfaces of the 11 components for which additional redesign data was collected. First, we test the null hypothesis that a change on interface (i,j) is independent of whether or not team i reports interaction with team j. To test that null hypothesis we categorize the 57 design interfaces according to two criteria:

- **First criterion**: whether or not the design interface *(i,j)* is matched by a team interaction reported by team *i*

- **Second criterion**: whether the design interface is the result of a change from component *i*

The results of the chi-square test of independence shown in Table B.1 do not allow us to reject the null hypothesis that.

Table B-2. Chi-square test results. Effects due to change of the interface

	Total	Expected number (fraction) of design interfaces matched by team interactions	Expected number (fraction) of design interfaces not matched by team interactions	Actual number (fraction) of design interfaces matched by team interactions	Actual number (fraction) of design interfaces not matched by team interactions	χ^2 of design interfaces matched by team interactions	χ^2 of design interfaces not matched by team interactions
Design interfaces with no change	39	24.632 (63.16%)	14.368 (36.84%)	25 (64.10%)	14 (35.90%)	0.005	0.009
Design interfaces with change	18	11.368 (63.16%)	6.632 (36.84%)	11 (61.11%)	7 (38.89%)	0.012	0.020
Total	57	36.000	21.000	36	21	**0.017**	**0.029**

Similarly, we test the null hypothesis that the impact on component *i* due to changes in interface (*i,j*) is independent of whether or not team i reports team interaction with team *j*. The results exhibited in Table B.2 do not allow us to reject such a null hypothesis.

Table B-1. Chi-square test result. Effects due to impact of the interface

	Total	Expected number (fraction) of design interfaces matched by team interactions	Expected number (fraction) of design interfaces not matched by team interactions	Actual number (fraction) of design interfaces matched by team interactions	Actual number (fraction) of design interfaces not matched by team interactions	χ^2 of design interfaces matched by team interactions	χ^2 of design interfaces not matched by team interactions
Design interfaces with no impact	43	27.158 (63.16%)	15.842 (36.84%)	27 (62.79%)	16 (37.21%)	0.001	0.002
Design interfaces with impact	14	8.842 (63.16%)	5.158 (36.84%)	9 (64.29%)	5 (35.71%)	0.003	0.005
Total	57	36.000	21.000	36	21	**0.004**	**0.007**

B.5. Sample of survey used

Review Redesign of Interfaces

System: ____________________________________

Representative: ______________________________

1) **What <u>portion of key interfaces</u> that your system has with each of the others was <u>changed by your system</u>?**

HPT Blades:	__None	__Less than half	__Half or more
HPT 1V:	__None	__Less than half	__Half or more
HPT 2V:	__None	__Less than half	__Half or more
HPT Rotor:	__None	__Less than half	__Half or more
HPT Case/OAS:	__None	__Less than half	__Half or more
LP Shaft:	__None	__Less than half	__Half or more
LPT Case:	__None	__Less than half	__Half or more
TEC:	__None	__Less than half	__Half or more
LPT Vanes:	__None	__Less than half	__Half or more
LPT Blades:	__None	__Less than half	__Half or more
LPT OAS/Tducts/Insulation:	__None	__Less than half	__Half or more

2) **What was the <u>impact to your system</u> caused by key interfaces changes made by each of the other systems?**

HPT Blades:	__None	__Minor	__Major
HPT 1V:	__None	__Minor	__Major
HPT 2V:	__None	__Minor	__Major
HPT Rotor:	__None	__Minor	__Major
HPT Case/OAS:	__None	__Minor	__Major
LP Shaft:	__None	__Minor	__Major
LPT Case:	__None	__Minor	__Major
TEC:	__None	__Minor	__Major
LPT Vanes:	__None	__Minor	__Major
LPT Blades:	__None	__Minor	__Major
LPT OAS/Tducts/Insulation:	__None	__Minor	__Major

C. An Algebraic Model to Address Various Task Assignments

In this appendix we describe an algebraic model that allows us to relax the assumption of one-to-one mapping between the architecture of the product and the development organizations. The model described here allows mapping a development organization split in *m* design teams to a product decomposed into *n* components. The model allows us to consider teams that participate in the design of various components as well as components designed by several teams.

C.1. Model Description

If component *i* depends on component *j*, then the teams designing those components are expected to interact to solve such a technical interface. In algebraic terms, such observation leads us to the following definition:

$$\mathbf{T} = \mathbf{C}^{T}\,\mathbf{A}\,\mathbf{C} \qquad \text{(C.1)}$$

where,

$\mathbf{T}_{m,m}$ = **Potential Team Interaction Matrix**. Matrix that contains the number of technical interfaces each pair of teams needs to resolve during the design process.

$\mathbf{t}_{i,j}$ = Number of interfaces that team *i* needs (potentially) to resolve with team *j* to complete its design(s).

$\mathbf{A}_{n,n}$ = **Binary Design Interface Matrix**.

$\mathbf{a}_{i,j}$ = 1, if component *i* depends on component *j* for functionality.

$\mathbf{a}_{i,j}$ = 0, otherwise.

$\mathbf{a}_{i,j}$ = 0, if $i=j$

For a development organization, arranged in "**m**" design teams, that designs "**n**" components, we define the following matrix:

$\mathbf{C}_{n,m}$ = **Design Contribution Matrix**. It is a binary matrix that contains whether or not a team directly contributes ("beyond the interface") to the design of a component.

$\mathbf{c}_{i,j}$ = 1, if team *j* directly contributes to the design of component *i*.

C.2. Proof of the Model

First, let us solve $\mathbf{D} = \mathbf{A}\ \mathbf{C}$, where row i of $\mathbf{A}$ contains the components on which component i depends on, and column j of $\mathbf{C}$ indicates which components are designed by team j. Hence,

$$\mathbf{d}_{ij} = \sum_{k=1}^{n} \mathbf{a}_{ik}\, \mathbf{c}_{kj} \qquad \text{(C.2)}$$

Note that $\mathbf{d}_{ij} \neq \mathbf{0}$ only for the cases where component i depends on component j AND component j is designed by team j. More specifically, $\mathbf{d}_{ij}$ equals the sum of interfaces of component i on which team j participates (as designer of other components).

Now, let us solve $\mathbf{T} = \mathbf{C}^{\mathbf{T}}\ \mathbf{D}$, where row i of $\mathbf{C}^{\mathbf{T}}$ indicates which components are designed by team i, and column j of $\mathbf{D}$ indicates the components that share interfaces with components designed by team j. Hence,

$$\mathbf{t}_{ij} = \sum_{k=1}^{n} \mathbf{c}^{\mathbf{T}}_{ik}\, \mathbf{d}_{kj} \qquad \text{(C.3)}$$

Hence,

$$\mathbf{t}_{ij} = \sum_{k=1}^{n} \mathbf{c}^{\mathbf{T}}_{ik} \left(\sum_{p=1}^{n} \mathbf{a}_{kp}\, \mathbf{c}_{pj} \right) \qquad \text{(C.4)}$$

Note that $\mathbf{t}_{ij} \neq \mathbf{0}$ only when team i designs a component(s) that depends on a component(s) designed by team j. Furthermore, $\mathbf{t}_{ij}$ equals the sum of interfaces of components designed by team i which depend on components designed by team j.

For the particular case when the mapping between components and teams is one-to-one (i.e. the design of each component is assigned to a design team) the Contribution Design Matrix ($\mathbf{C}$) becomes the Identity matrix. Hence,

$$\mathbf{T} = \mathbf{A} \qquad \text{(C.5)}$$

C.3. Properties of the Model

The model specified in (C.1) exhibits several properties derived from linear algebra theory (Strang, 1980).

C.3.1. Associative property

$$\mathbf{T} = \mathbf{C}^T (\mathbf{A}\ \mathbf{C}) = (\mathbf{C}^T \mathbf{A})\ \mathbf{C} \qquad \text{(C.6)}$$

C.3.2. About symmetry of T

If **A** is symmetric, then **T** is symmetric.

C.3.3. *Inertia Law* of Sylvester

If **C** is square and no singular matrix (i.e. **det** $\mathbf{C} \neq 0$), then $\mathbf{T} = \mathbf{C}^T \mathbf{A}\ \mathbf{C}$, has the same number of positive eigenvalues than **A**, the same number of negative eigenvalues and the same number of zero eigenvalues. Therefore, the signs of the eigenvalues of **A**, but not their values, are preserved.

C.3.4. Markov transition matrix

T could represent the transition matrix of a Markov process associated to the interactions between design teams if the rows of **T** sum less than 1 and all the entries of **T** are non-negative.

D. From Printouts to Parameters (log-linear model)

The values found in the printouts provided by commercial statistical packages, such as GLIM, BMDP, SYSTAT, or SPSS, need to be translated to obtain estimates of parameters of the p_l-based models (for details refer to Wasserman and Faust, 1994 pp. 665, and Wasserman and Weaver, 1985). Since we used SPSS (version 10.0) to complete our analysis, we illustrate the translation process for αs and ρ of p_l given the printouts provided by SPSS 10.0.

Note that the translation process described here is different than the description offered by Wasserman and Faust, 1994. The difference is due to the fact that SPSS (newer versions) uses the constraint that the last row and the last column of any set of u-terms are defined to be zero (they are called "aliased" or redundant parameters). For example, the estimated u-terms produced by SPSS corresponding to the r x k_l interaction terms of model 1 of Table 5.2 are the following:

	K1=0	K1=1
R=1	-0.1452	0
R=2	0.0779	0
R=3	0.2989	0
R=4	0.3041	0
R=5	0.8522	0
R=6	0.364	0
R=7	0.5436	0
R=8	0.0000	0

To translate the r x k_l u-terms to α_ls, we first need to "recenter" the first column by computing its mean and subtracting it from each of its elements. That is,

	K1=0	K1=1
R=1	-0.1452- 0.28694 = -0.43214	0
R=2	0.0779 - 0.28694 = -0.20904	0
R=3	0.2989 - 0.28694 = 0.012	0
R=4	0.3041 - 0.28694 = 0.0172	0
R=5	0.8522 - 0.28694 = 0.5653	0
R=6	0.364 - 0.28694 = 0.0771	0
R=7	0.5436 - 0.28694 = 0.2567	0
R=8	0.0000 - 0.28694 = -0.28694	0

The α_l parameters are obtained by multiplying the values of the first column by (-1), which correspond with the values of α_l (α parameters of the design interface matrix of model 1) reported in Table 5.2. We proceed similarly with the r x l_l u-terms to obtain the values of the β parameters.

Since the constraints used by SPSS are the same we use for the ρ parameter, the u-term k_l x l_l corresponds to the estimate of ρ_l, no further adjustment is required. Standard errors of the model parameters estimates can be also derived from slight adjustments of the standard errors of the u-terms (for details see Wasserman and Weaver, 1985).