Technology for a Smarter Planet: The role of Cognitive Technologies and Open Innovation

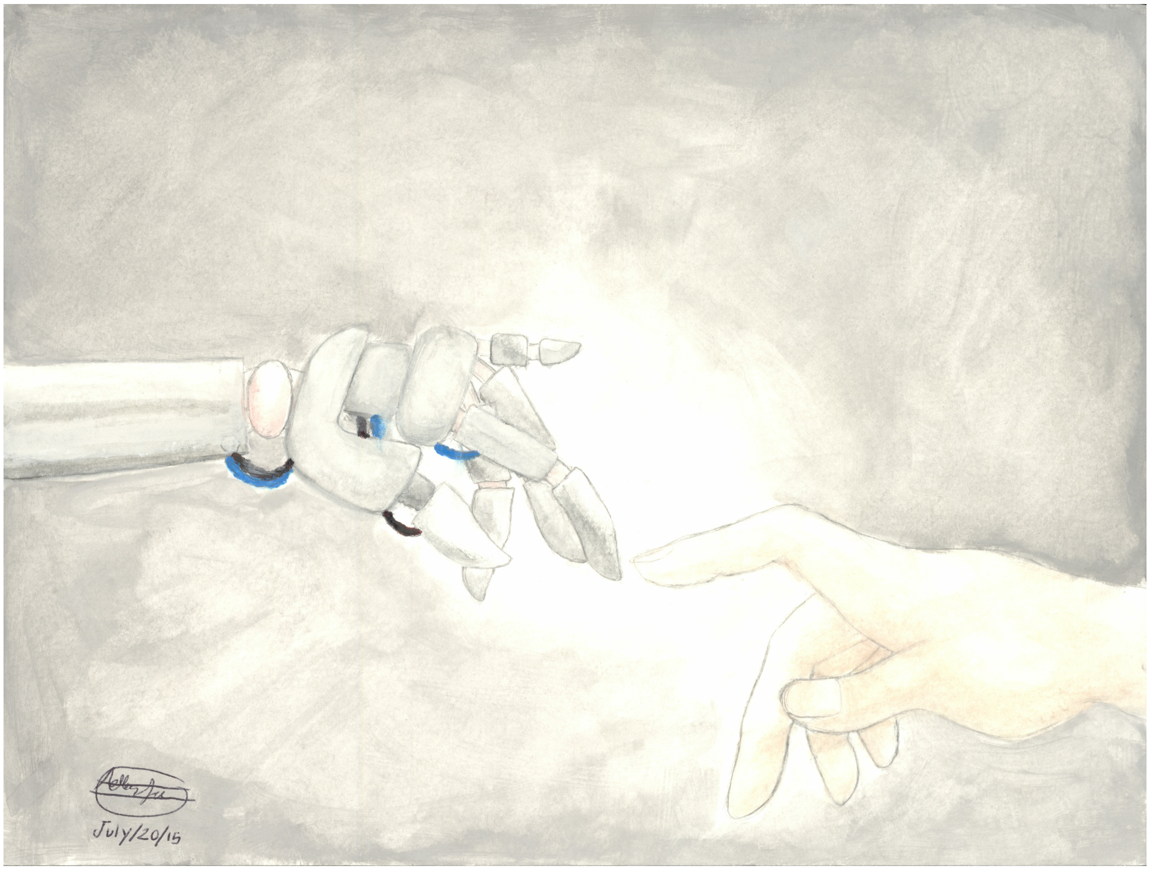
by

Ajit Sharma

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in The University of Michigan
2015

Doctoral Committee:

Professor Mayuram S. Krishnan, Chair
Professor Gautam Ahuja
Professor Robert J. Franzese Jr.
Assistant Professor Yan Huang


July/20/15

To the memory of my grandfather, Dr. Shatrughan Prasad Sinha (1926-2007),

freedom fighter, idealist, and a man whose life was his message.

For my parents.

ACKNOWLEDGEMENTS

My love affair with Ann Arbor and association with the Ross School of Business goes back to when it was simply the University of Michigan Business school and I was an evening MBA student. Comparing my MBA and PhD journeys, I am amazed at how much is given unconditionally to doctoral students at Ross: unfettered access to esteemed scholars, financial resources, 24X7 access to an office space, a supportive doctoral office; the support of dedicated cleaning, facilities, faculty support and technology staff. The extended support system in the wider university is equally impressive - 24X7 access to peerless libraries and study spaces; shuttle service till 3 am; the Michigan Union for a quick snack or 40 winks on a couch in the quiet study hall; an embarrassment of eateries on state street that serve food from across the world; the list goes on...I am indebted to the University of Michigan and the Neverland called Ann Arbor, if ever there was one.

Looking back at my PhD journey, the guidance and support of my advisor, Dr. M.S. Krishnan was invaluable. Like a master coach, he nudged, pushed, critiqued, challenged and encouraged me as the situation demanded. Under his tutelage I was extended to discover my limitations and strengths. On the other hand, he was

willing to listen me out and offer his advise as a friend, when I needed it. His unassuming and accessible nature belie his stature in the IS research community and reflect his commitment to students. I especially wish to thank him for giving off generously from his time in spite of the multiple demands of associate deanship, executive education, MBA classes and building up Ross school's global footprint. In the instances when I saw him juggle the many hats he wears, I was humbled by his capacity to dispense each of these roles simultaneously and flawlessly. His strong work ethic, the desire to contribute to the school, and maintaining an always smiling demeanor in the midst of so much activity are some of his many traits that will stay as aspirational states for me. On the research front, I've benefitted immensely from his presence and advise. His ability to review a draft, hone in on its gaps and identify the needed improvements, all within a matter of minutes saved me iterations that would have been needed otherwise. His clear-eyed advise on how to manage the duality of exploration versus exploitation in research, has been crucial in achieving closure on projects. I owe to his mentoring my growth as a researcher as well as a person. I've been fortunate to know him as a teacher, scholar and human being and the example he has set will stay with me.

Dr. Yan Huang joined Ross towards the second half of my program. In addition to mentoring me she provided encouragement like a friend. Her knack for focusing on my strengths meant that I always left her office with lifted spirits. Even though Dr. Yan is a star among young scholars in our field, she always came across as modest and accessible. She made herself available despite her busy schedule. She painstakingly reviewed drafts and provided meticulous feedback above and beyond

the call. I must also thank Dr. Yan for helping me practice talks and interviews. She has been part advisor and part friend throughout my journey. I'm indebted to her for the unqualified and abundant support.

My first introduction to Dr. Ahuja as an MBA student was as one of the most sought after professors and had the good fortune of getting a spot in his overbooked classes. As a PhD student I witnessed a different side of him as a scholar of world renown. He belongs to a select group of professors who have excelled both at teaching and research. My work has benefited immensely from his mastery of the strategy field which he willingly dispensed to enrich my research with perspectives outside the IS domain. I am also grateful to him for accommodating demands on his time during a very busy sabbatical year at Harvard. Dr. Rob Franzese, as the outside member on my committee helped ensure the methodological rigor of my thesis. I am especially thankful to him for being immediately available despite his hectic schedule. I feel blessed to have had Dr. Krishnan, Dr. Yan, Dr. Ahuja and Dr. Franzese as my committee members.

Apart from the members of my committee, I owe a debt of gratitude to many professors at Ross. First of all, I am grateful to Dr. Nigel Melville for the time and energy he invested in his seminars. Especially, his seminar on the Theory of IS has been foundational in helping me develop some understanding of the IS field. As the doctoral program coordinator, Dr. Michael Gordon was extremely helpful in easing us fresh doctoral students into the program. I enjoyed and benefitted greatly from his seminars. His seminar on the nature of technology provided a means to view technology with new eyes. Similarly, Dr. Scott Moore's seminar on the philosophy of science provided a lens that seems to keep giving. Dr. Hila Etzion's seminar was essential in rounding off the methods toolkit.

Seminars from Professors from other disciplines such as Dr. Karl Wieck, Dr. Gautam Ahuja, Dr. Ravi Anupindi, Dr. Jerry Davis, Dr. Sendil Ethiaraj and Dr. Jun Li, have benefited me in improving my abilities. I'm also thankful to Dr. Amitabh Sinha for being a patient header through the program. Dr. Brian jones was naturally given to helping out like a friend and providing sage advice drawing on his doctoral experiences. I also owe a debt to professors from the schools of information, computer science, public policy and education for their courses.

I'm grateful to my cohort Dan Rush and other members of the BIT group who made for a caring and supportive community - Terence Saldanha, Suresh Malladi, Andrea Walrath, Min-Seok Pang, Sanghee Lim and Mark Madrilejo. I'd especially like to thank Terence for being a sounding board, a guide and dear friend during our time together at Ross as well as after. Suresh, for the many discussions we had late at night. Rahul Chabra for the discussions about the rhythms of an academic life. I'd also like to thank my officemates in R3431 for the years spent together - Gareth Keeves, Maddy Ong, Jin Woo, Eun-Soo and Guy Shani. I'd especially like to thank Gareth for the many hours spent happily typing away in R3431, for the stimulating conversations and the sharing of his polymathic knowledge across disciplines. I'd also like to thank my many PhD co-mates for their friendship and companionship including Sutirtha Bagachi, Mike Palazolo, Jihoon Cho, Vivek Tandon, Pranav Garg, Shripad Devalkar, and Adithya Pattabhi.

Finally, I wish to remember the person who inspired me to quit industry and enter academia - the late Prof. C.K. Prahalad. I had the good fortune of working with him on his research on the Bottom of the Pyramid and remember with fondness the energy and passion with which he filled up any room he stood in. I also wish to thank Prof. Gajanan Kulkarni, Prof. Subhash Chandra Rastogi and Prov. Harsh V. Bhasin for encouraging me to take this path.

For PhD students with a family, it may not be an exaggeration that it is undoable without family support. And yet, I feel especially lucky for the faith my family placed in me and the unconditional flexibility to do what I wished to do with my time. I pursued my PhD program from a home 60 miles away from campus. I'd complete freedom to leave home at any hour without notice, be gone for days on end only to return to welcoming arms, good food and a loving family. This, despite the many transgressions of missing birthdays and other days worth cherishing. It is at these moments that I was reminded that love any day beats a life of the mind. I could not have done this without you Pratibha, Aditi and Anjali! Thank you for this gift.

My parents were a constant anchor and nourished my spirit. Like clockwork they kept calling from India to check up on me and ensure that I and the family were doing well. My mom always had a quote or two from the Geeta and kept reminding me that one who takes care of all will see us through. She took deep interest in my research, listened to my wild ideas, and provided her own unique take on what seemed interesting and what was banal. My dad provided constant encouragement and believed in me. His occasional queries on when I was planning to be done were enough to double up my efforts. My brother Amit and sisters Bharati, and Rashmi were a constant source of moral support.

I thank my two girls, Aditi and Anjali, for being a reminder that I could not rest. There was a sharp jump in my productivity after I shared my deadlines with them and they started to check up on my progress at least as many times as I inquired about their homework. I am grateful for their forbearance as they watched me downgrade their lifestyle from that of a double income household to that of a resource constrained one. My comeback has to be that adversity is a better teacher than affluence. To the little one on her agonizingly simple yet unanswerable question...papa why are you still in school? Darling, guess what? I'm done.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

<u>**Table**</u>

ABSTRACT

A Smarter Planet: The role of Cognitive Technologies and Open Innovation

by

Ajit Sharma

Chair: Prof. M.S. Krishnan

Information technology is making the human race smarter by increasing its cognitive capacity through at least two drivers: Open Innovation and cognitive computing. Open innovation allows leveraging the wisdom of the crowds by bringing in more people into the fold through open innovation platforms, open source development and citizen science. In this sense, open innovation is enabling harnessing of the latent cognitive surplus of the human race. The dawn of the cognitive computing era on the other hand, is affording new uses of computers in organizational decision making. Specifically, IT is now enabling organizations to scan, interpret and learn from larger subsets of their informational environment hitherto considered inaccessible and uninterpretable by computers. As organizations and individuals gain this sixth sense of sorts, they can make better resource allocation decisions.

In my dissertation, I study both these technology developments and their role in

making organizations smarter and thus better generators of value. The underlying motivation is that better organizational decisions will allow better use of scarcer resources making the planet more sustainable.

In chapter 2, a purposive theoretical framework for synthesizing the role of IT in organizational decision making is attempted. The proposed *the interpretive model of IT*, also achieves a clear delineation between the programmable and cognitive computing eras. In chapter 3, the antecedents of predictive analytics usage within firms are explored through an empirical study. In chapter 4, an empirical study of the idea selection process within an open innovation funnel is undertaken to explore the question whether open innovation funnels prefer innovative ideas over conservative ideas.

CHAPTER I

Introduction

1.1 Introduction

1.1.1 Research Motivation: Spaceship Earth and its limited carrying capacity

The characterization of Earth as a spaceship hurtling through space [Boulding, 1996], gains increasing relevance with the growing realization that Earth is neither a wellspring of limitless resources nor an infinite cesspool. With time, the carrying capacity of this spaceship is likely to be tested to its limits in light of its growing crew size and limited resources. Arguably, in our era, the principal agent making decisions for allocating these scarce resources to purposive uses is the *organization* (for-profit, Governmental and non-profit). Extending Boulding's analogy, if we consider organizations as the pilots steering spaceship earth, till recently, the dominant paradigm has been that of a closed cockpit door with pilots inside doing the flying. Two recent developments challenging these assumptions are radical openness and cognitive computing. First, with radical openness, the cockpit door is opening up slowly and the wisdom of the crowds is being harnessed. Second, with the dawn of the cognitive computing era, as computers acquire interpretive abilities, machines are graduating from *aiding* to *augmenting* human decision making.

In my dissertation, I contend that as organizations look for ways to push the frontiers of their value creation potential in their quest to do more with scarcer resources, both the shift towards radical openness as well as cognitive computing will become indispensable allies and permeate more aspects of the organizational form and function.

Within the wider scope of cognitive computing, I explore the role of advanced analytics in helping organizations make better decisions. Advanced analytics should allow organizations to harness the intelligence of machines to overcome the bounded rationality of the behavioral economists predictably irrational man. Just as machines flying planes has reduced instances of grounded planes, machine augmentation should reduce instances of organizational decisions that destroy value. Within the wider scope of radical openness, I specifically explore the role of open innovation in enabling organizations to harness expertise from outside the firm in their quest to become more innovative. As technology enables organizations to tap into the global nervous system, it is likely to benefit firms that are able to harness the hitherto unharnessed pool of human ingenuity.

In sum, my thesis is focused on the enabling role of technology in helping organizations make better decisions both in managing innovations as well as ongoing operations (Figure 1.1).

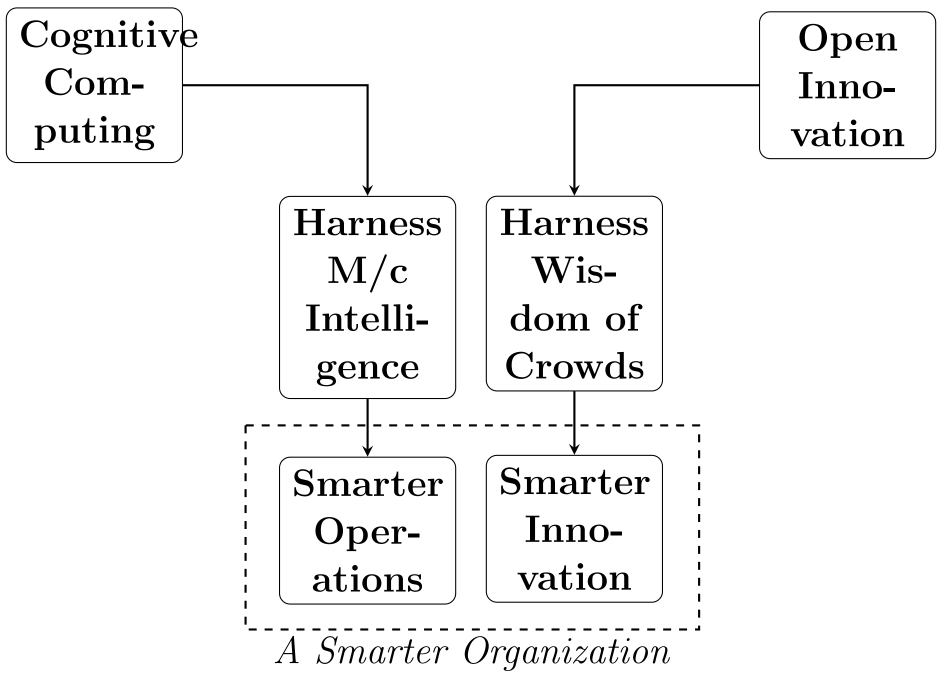


Figure 1.1: Thesis Overview

I start out in chapter II with an extensive review of IS literature related to decision making. This review reveals that IS literature does not seem to have a theoretical framework which can accommodate the new role of IT as a decision maker. Thus, literature in other areas such as sense making, decision making and bounded rationality is drawn upon to develop a vocabulary and theoretical framework for working with the cognitive computing era. Specifically Daft and Wieck's interpretive model of organizations is adapted to develop the interpretive model of IT. In developing this new framework, in effect, I propose to shift from the widely accepted 'Automate-Informate-Transformate' model of the role of IT to an 'Automate-Informate-Transformate-*Interpret-Learn*' model of IT. From a practitioners viewpoint, a Quality Function Deployment framework is developed to illustrate the usefulness of this model in helping organizations map IT capabilities to desired organizational outcomes.

The proposed interpretive model of IT provides an alternative to several accepted modes of considering the IT construct within IS literature. First, it departs from the traditional approach of black boxing the IT construct to prying open its innards to identify its components. Second, it attempts to go beyond earlier attempts to deconstruct IT as an ensemble of technologies or capabilities to considering it in purposive terms. Specifically, it extends the Automate-Informate-Transformate model of IT [Zuboff, 1988]. Finally, from a practitioners point of view, it provides an alternative framework to technology hype cycles as the dominant lens through which IT investments are considered by firms.

In Chapter III, the proposed interpretive model of IT thus developed in Chapter II, is used to clearly distinguish between traditional and advanced analytics. Additionally, a taxonomy of organizational interpretive modes relevant in the cognitive computing era is developed. This theorizing is leveraged to empirically test hypothesis about the antecedents of firms that adopt advanced analytics versus those that remain with traditional analytics.

While the subject of organizational decision making, which has been the focus till chapter III, lies squarely within the skin of organizations, the phenomenon of open innovation, studied in chapter IV, exists in markets that mediate transactions between organizations as well as between organizations and individuals. Indeed, the recent cross-category shift towards radical openness, enabled by technology, is possibly redefining what gets done within organizations versus in markets. The focus of chapter IV is open innovation. Specifically, we empirically study an open innovation contest in a large private bank to explore whether it encourages the firm to select more innovative ideas.

CHAPTER II

Interpretive Model of IT: Towards a Theory for the Cognitive Computing Era

2.1 Decision Making in Organizations

Administration is ordinarily discussed as the art of "getting things done." Most classical analyses of organizations have emphasized the division of work as the basic characteristic of organized activity. Gulick, for example, in his "Notes on the theory of organization," says: "Work is the foundation of organization: indeed, the reason for organization." Other classical paradigms that have dominated modern day management are Taylor's focus on efficiency and Fayol's division of labor. Not very much attention has been paid to the choice that precedes all action - to the determining of what is to be done rather than to the actual doing. However, any practical activity involves both "deciding" and "doing". In fact, the task of "deciding" pervades the entire administrative organization quite as much as does the task of "doing". Indeed it is integrally tied with the latter.

The relative lack of focus on decision making in classical and neoclassical organization literature can possibly be attributed to the underlying assumption of the rational economic man. Much of management literature has been prefaced on the assumption that human beings are perfectly rational agents trying to maximize a known utility function. However, this unqualified belief in the perfect rationality of human decision making has been called into question by the introduction of the notion of "bounded rationality" by Herbert A Simon. Most notably, Herbert Simon created a significant body of work which emphasized how organizations can be understood in terms of their decision processes.

> *A general theory of administration must include principles of organization that will insure correct decision-making, just as it must include principles that will insure effective action...Even though, as far as physical cause and effect are concerned, it is the machine gunner and not the major who fights battles, the major is likely to have a greater influence upon the outcome of a battle than any single machine gunner.*
>
> *- Administrative Behavior, Herbert A. Simon, 1945*

It is the purposes of this thesis to explore this centrality of decision making in organizations. Specifically, I study how advances in technology are enabling it to play an increasingly important role in organizational decision making.

2.2 Decision Making in Organizations: Changing Role of Technology

The bias towards execution in organization theory is also reflected in IS literature and in IT practice. Much of the focus in IT, till recently, has been on the use of computers to perform more rapidly and cheaply than before the same functions that we formerly carried out with adding machines and type-writers. Most IT projects have been variants of business process reengineering which are focused on eliminating inefficiencies by automating or transforming business processes. At its core these process improvement projects use fixed procedural logic to automate the execution of tasks and work flows. Indeed, the selling proposition of many Commercially Off The Shelf (COTS) enterprise systems such as ERPs and CRMs is the encapsulation of best in class processes in software form. Thus, standardized processes such as accounting, distribution, shop floor control and customer relationship management can be immediately automated in a firm by the purchase and implementation of enterprise systems. These systems have also been grouped under the moniker of Online Transaction Processing (OLTP) systems since they record transactions which serve as the system of record for the tasks and activities that make up the day to day operations of a firm.

Online Analytical Processing (OLAP) systems rather than generating transactions, generate reports and visualizations on historical transactions generated by OLTP systems. Though their characterization as Business Intelligence (BI) software might seem to allude to an ability of the software to provide intelligence, they simply fetch historical transactions and present them as reports and increasingly as advanced visualizations. The actual interpretation of the information and decision making happens in the minds of a human being. In that sense, BI software is an aid for human decision making rather than an augmenter.

Needless to say, there has been a long interest in creating computing machines that can think and make decisions on their own [**?**]. However, apart from some areas of middle-management decision, where techniques like linear programming (from operations research) and expert systems (from artificial intelligence) are now widely employed, computers have changed executive decision-making processes only modestly. However, technological advancements in the recent past and specifically the past few years, are redefining the role of IT in organizational decision making. In the next few sections, I briefly discuss some of these developments.

2.2.1 Pervasive digitization and the scarcity of attention in an information-rich world

A recent technological trend is the pervasive digitization of all aspects of an organization and its environment. This has resulted in an orders of magnitude expansion of the informational environment of a firm which can be represented digitally. This pervasive digitization is leading to the era of big data. It is estimated that the data volume in the world is doubling every year. Indeed we have squarely shifted from the age of data poverty to the era of data glut. Organizations are being overwhelmed and are grappling to manage this data deluge. In addition to volume of data, its velocity and veracity is also a significant challenge in making sense of a firm's big data assets. Thus pervasive digitization is compounding the problem of bounded rationality that plagues human decision making. Whereas technology has engendered a situation wherein the bounded rationality of humans is being exceeded, paradoxically technology developments in a different quarter (Cognitive computing or Data science) is coming to the rescue by augmenting human decision making.

2.2.2 Dawn of the cognitive computing era: A technology side-kick for organizational decision making?

There has been a long history of interest in designing computer systems that can think like humans. Expert systems and Decision Support Systems (DSS) are examples of such programs that have been developed and employed by firms. An expert system usually consists of several parts: (1) a knowledge base, where the expert knowledge resides; (2) a database, where historical and new data are stored; (3) an inference engine, which provides different types of inferences; (4) an efficient interface to users; (5) an explanation module, which provides an explanation of how and why a certain decision was recommended by the system; and (6) a module that learns and accumulates new knowledge, based on the systems operation and on new incoming data [**?**]. In the 1980s, several firms experimented with such systems, however due to limited success, interest plummeted by the late 1980s. Decision support systems as the name suggests, help humans make a decision about a given problem, under given circumstances and constraints. In contrast *Decision making systems* are systems that make final decisions and potentially take actions. Some examples are automated trading systems on the Internet, systems that grant loans through electronic submissions and email spam filters. The technology that underpins these cognitive technologies are algorithms such as machine learning, neural networks, pattern recognition, image processing and natural language processing. There has been an upsurge in applications of cognitive technologies in a wide range of areas such as business, medicine, cyber security etc. This jump from decision support systems to decision making systems has been heralded by IBM as the dawn of the cognitive computing era. Indeed, to the scholars and practitioners who witnessed the exuberance around DSS and Expert Systems in the 80s, this enthusiasm for cognitive technologies might seem like *deja vu* all over again. However, I contend that there are important differences which might mean that the cognitive technologies will become more pervasive and successful that Expert system. One of these is that the cognitive technologies are generic and modular and do not need to built from the ground up for each specific application. This is discussed below.

2.2.3 Building Blocks of Cognitive Technologies: Generic, Modular and Widely Applicable

Expert Systems and Decision Support Systems, are restricted by their architecture to be applicable to very specific and narrow domains. The knowledge of human experts in a particular field is extracted through a methodology called Knowledge Engineering and this knowledge is then captured in the 'rules engine' of the software. Thus, we have expert systems for diagnosing the fault in an engine, for selecting the right material for a particular product, for diagnosing the possible root cause of defects in manufactured products etc. In effect, Expert Systems are tightly wedded to a specific application. Indeed, one of the highly valued feature of knowledge based systems is the explanation component which can be requested to explain the reasoning followed by the software to reach its conclusions[Gregor and Benbasat, 1999]. Advanced analytics applications on the other hand leverage generic algorithms which are being made available increasingly as add on packages of statistical tools such as SAS and MATLAB. The implication of this difference is that as libraries of machine intelligence algorithms become part of the core of generic statistical packages, it should be possible for domain experts in different domains to build applications leveraging these algorithms. This difference should in turn reduce the expertise needed for building these applications. The implication of this is that the cognitive technologies can be made available as generic algorithms that can be employed for many different purposes. A knowledge of writing the algorithm is not required for a user to be able to make use of it.

2.2.4 Analytics as a Service: Democratic, Affordable, Pay Per Use Access to the cutting edge

A second implication of the generality of cognitive technologies is that it is possible to offer them as a service on a pay per use basis. Indeed there has been a proliferation of firms that offer analytics as a service over the cloud. Amazon Web Services and Splunk are instances of firms that offer analytics as a service. There was a time in the not too distant past when only firms above a certain scale could afford the latest technologies. This state of affairs was perpetuated by software ownership as the dominant mode of having access to computing power and software functionality. This meant that only large firms with deep pockets and the scale to leverage the complete set of bundled features of enterprise scale applications could benefit from it. However, with the utility model of offering computing, memory, software, analytics on the cloud, most IT functionality has been deconstructed so that firms of any scale can choose to use it as a utility. The implication of this development is that machine intelligence is within reach of organizations with a range of scale.

2.2.5 Accelerated Innovation in the App Economy: From Biological to Artificial Evolution

The software industry seems to have matured a model of innovation that serves as an alternative to closed door innovation within R&D centers of large firms. This is known as ecosystem innovation ([**?**]) and in the software world, it is also sometimes referred to as the App economy. In the app economy there is a division of labor between the platform owner and app developers. Platform owners focus on operationalizing a development platform which is powerful as well as easy to use for building apps. App developers on the other hand focus on building apps. This achieves a division of labor in the innovation process. In effect, the innovation that end users consume is being executed by app developers at the tip of the technology stack. The app developers do not expend their energy and attention in learning, developing and maintaining the full technology stack. The platform orchestrators on the other hand need focus only on developing the core underlying platform technology and the tools for the app developers that are powerful and easy to work with.

The platform also serves as a distribution channel through which end users consume the apps. End users serve as the evaluation mechanism which judges the 'fitness" of the apps. Apps are voted up or out by the likes and wallets of end users. The uncanny resemblance of ecosystem innovation to evolution is unmistakable. What is different is that unlike biological evolution, variation and selection occurs at a very high clock speed. For illustrative purposes, Apple's app marketplace is a very efficient market mechanism for bubbling up to the best applications that customers have voted with their likes and their wallets.

This platform centric innovation which started in the consumer space in the form of platforms such as Apple and Android are now diffusing into enterprise scale systems such as SAP's Hana platform, Salesforce.com's platform. In the field of cognitive technologies, IBM has opened up its Watson technology as part of its BlueMix platform to anyone interested in building apps leveraging the same cutting edge technology that defeated the world chess and Jeopardy champions. This is a clear illustration of how ecosystem innovation is a force multiplier in innovation by segregating the innovation between orchestrating firms and participating partners. For any firm to replicate the cognitive technologies embedded in IBM Watson would be a formidable task. The implication of this model of innovation is that if executed well, the clock-speed of innovation is going to increase by orders of magnitude thus resulting in a continuous improvement in the applications of cognitive technologies available for end users.

Domain Agnostic Innovation in the Ecosystems: Variety and Relevance A second implication of ecosystem innovation is that the app developers from widely different domains could be using the same underlying technologies to build apps relevant to their domain. Thus the Watson platform could be used by developers building apps in fields of demand forecasting; jet engine failure prediction; cancer detection; identifying fashion trends; fraud detection etc.The apps are likely to be developed by development firms that are closest to the domain for which they are building their apps. and hence are likely to be most relevant and feature rich. In effect, the domain knowledge portfolio of the orchestrating firm is not the bottleneck for the scope and pace of innovation. The implication of platform centric innovation is that the cognitive technologies are unlikely to remain esoteric experiments within the closed doors of organizations with large R&D budgets. Rather, we are likely to see a cross-category proliferation of applications the best of which should become available for all to use.

2.2.6 A perfect storm for a new role for IT in decision making

The above discussions reveal a dual push of technology on organizational decision making. On the one hand, pervasive digitization has brought in the era of data glut aggravating the bounded rationality of human decision making. On the other hand, the availability of cognitive technologies and advanced analytics as a service and as apps implies that even small firms are likely to have access to a sixth sense of sorts that can help them leverage their big data assets to make better decisions.

Though there is extensive coverage in the popular press of the transformational effect of the cognitive computing era on organizations , the extent of its treatment in the IS scholarly literature is not clear. Thus, we next conduct an extensive review of the IS literature to ascertain how these trends have been addressed in IS literature.

2.3 Decision making in IS literature: A Review

The ProQuest database was searched for papers in MISQ and ISR journals based on three different subject terms - Decision, Intellig* and Optimiz*. The number of articles returned for each of the subject terms searched for is presented in Table 2.1.

Table 2.1: Literature Review Article Count

Search Term	MISQ	ISR
Decision	101	62
Intellig*	17	11
Optimiz*	1	5

Some of the articles were about a specific decision such as inventory decision or IT investment decisions. Given the focus of my research, such papers were excluded from our consideration set. Only papers that were about the role of IT as a decision aid or augmenter of decisions were retained. Each article was then reviewed and categorized based on the type of article (Empirical, Case Study, Experimental, Design & Recommendation, Theoretical and Review), key artifact(Eg: Expert systems, Decision Support Systems, Distributed Decision Making System, Electronic Meeting System, Creativity support software, Algorithm etc.); and focal theme (Eg: System Design, System Performance, System Adoption, System Use, Business Value etc.).

A list of these papers with these codings is available in Appendix . A review of this literature reveals some key insights which are discussed next.

2.3.1 A history of interest in intelligent machines

A review of these papers reveals that a preoccupation with intelligent computers is by no means a recent phenomenon in the IS discipline [Kugel, 1988]. Indeed the 1980s saw an upsurge of interest in IS literature on expert systems and decision support systems. However, by the late 1980s that interest gradually faded away. Some of the factors identified for this drop in interest were lack of system acceptance by users, inability to retain developers, shifts in organizational priorities and problems in transitioning from development to maintenance [Gill, 1995].

2.3.2 Key Research Themes in IS Literature on Decision Making

As is made apparent by the thematic categorization of the articles in the Appendix, there are some clear themes that emerge in the literature. Some of the broad themes are reviews, frameworks, methods & tools, specific applications, adoption & usage, business value and design. Additionally, the technological artifact enabling organization decision making can broadly be classified into DSS and analytics. The literature on DSS seems to be concentrated in the 1980s. A recent trend of the past few years is an upswing in articles on the applications of AI centric analytics in the business domain. The applications of AI discussed in the literature range on a continuum from the general[Meyer et al., 2014] to the very specific. Some applications specific to a domain are meta-learning framework for detecting financial fraud

[Abbasi et al., 2012]; Identifying systemic risk in banking systems[Hu et al., 2012]; Business intelligence in blogs [Chau and Xu, 2012]; From business intelligence to competitive intelligence [Zheng et al., 2012]; Designing intelligent software agents for auctions with limited information feedback [Adomavicius et al., 2009]; Social network based inference model for validating customer profile data [Park et al., 2012].

2.3.3 A Lack of Theoretical Frameworks

Though our review reveled a long history of IS literature on decision making, we did not find any theoretical frameworks that attempt a synthesis or a theoretical framework. The absence of theoretical frameworks in our literature review is hardly surprising given the a-theoretic character of IS research[Orlikowski and Iacono, 2001]. The theoretical framing that seemed closest to our purpose is the [Zuboff, 1988] model of IT which explicates the different roles of Information Technology in an organization. However, though this framing is sufficient for the traditional role of IT is not sufficient for explaining the new role of IT as a decision augmenter in the cognitive computing era. Thus, a search beyond IS literature in other reference disciplines was conducted in an effort to construct a theoretical framework capable of accommodating this new role of technology for decision making. This endeavor is discussed in the next section.

2.4 Interpretive Model of IT: Towards a New Theoretical Framework For the Cognitive Computing Era

Our theoretical framework is inspired by the literature on sense-making[Weick, 1995a], organizational and managerial interpretation[Daft and Weick, 1984] and organizational and managerial cognition[Walsh, 1995]. Specifically, we draw heavily upon Daft and Wieck's interpretive model of the firm[Daft and Weick, 1984]. In building a model of organizations as interpretation systems, Daft and Weick(1984) note that:

> *The critical issue for interpretation systems is to differentiate into highly specialized information receptors that interact with the environment. Information about the external world must be obtained, filtered, and processed into a central nervous system of sorts, in which choices are made.*

In their interpretive model of the firm, they identify three stages that constitute the overall organizational learning process: Scanning, Interpretation and Learning (Figure 4.1). Scanning is defined as the process of monitoring and collecting data from the environment. This data collection can be through IT systems or human connections. Interpretation is the step where the data is given meaning. Based on the interpretation, action is taken and observations of outcomes generate learning.

We apply this interpretive model recursively at two levels - the IT system level and the organizational super-system level within which IT is situated. In the current

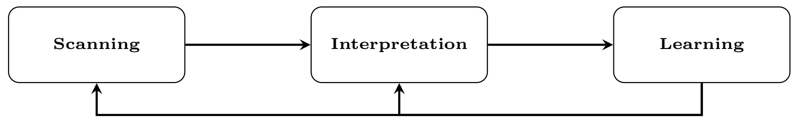


Figure 2.1: Organizations as Interpretation Systems
(Source:Daft and Wieck, 1984)

section, we apply the interpretive model to the IT system to arrive at its purposive components. In the next section we apply the interpretive model at the organizational level to build our theory and hypothesis.

Casting IT into Daft and Wieck's interpretive mold reveals a natural congruence between the services it renders for an organization and the three stages of scanning, interpretation and learning. We delineate some of these stages into sub-parts to better represent the division of labor in IT which in the human and organizational mind are arguably intertwined. We decompose the scan stage into scan and store; interpret stage into present and interpret; and the learn stage into execute and learn. An illustration of our proposed *interpretive model of IT* is presented in Figure 2.2.

To our knowledge, this is one of few attempts within the IS literature to theorize[Weick, 1995b] the IT artifact[Orlikowski and Iacono, 2001]. In an extensive review of IS literature, Orlikowski and Iacono found that in a significant percentage of IS literature, IT is in essence the *omitted* variable, nominally referred in passing, used to build context and routinely treated as a monolith. We believe that such 'black-boxing" of the IT artifact will become increasingly untenable given a need to understand the paradigmatic shift from the programmable to the cognitive computing era. The sense-making of the changes underlying this shift will require the IS

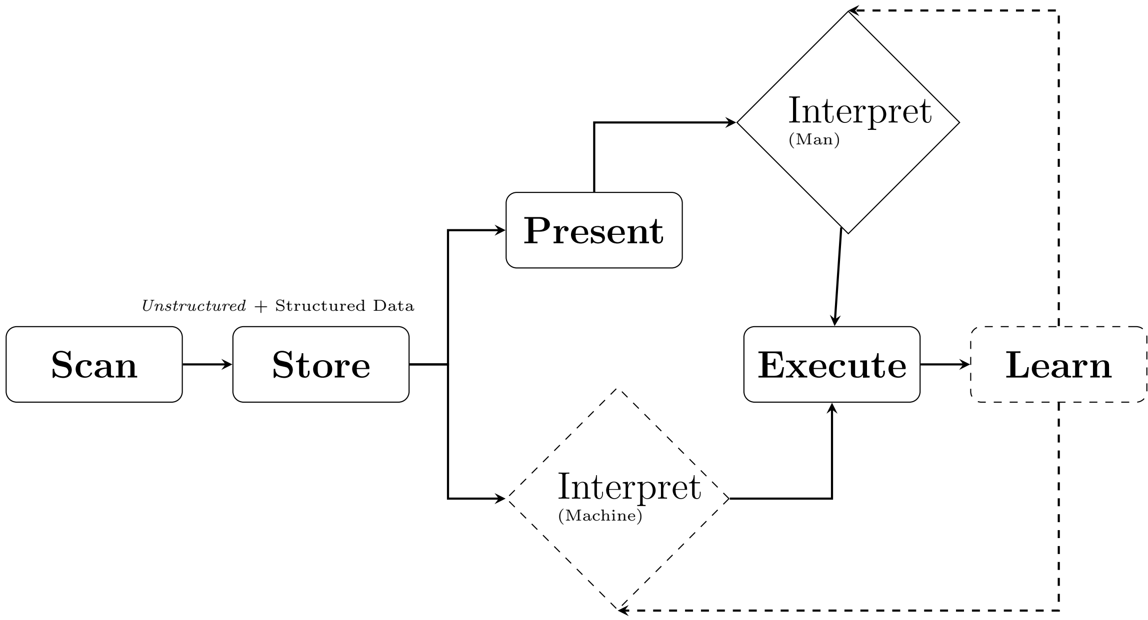


Figure 2.2: Interpretive Model of Information Technology
(Adapted from Daft and Wieck, 1984)

community to pry open the IT construct and peer inside to parse the components that are new, those that are changing and those that are becoming obsolete. We feel that in responding to Orlikowski and Iacono's call for theorizing the IT artifact, we have been amply rewarded with a framework which can be gainfully employed to deepen and advance our understanding of the IS field.

First, our model is a first attempt, to our knowledge, in IS literature to sharply delineate the distinguishing features of the cognitive computing era(dashed elements) from those of the programmable computing era(solid line elements).

Second, it serves as a possible taxonomy for organizing extant IS research. Some reflection on the model yields that much of extant IS research has been focused on the box labeled 'Execute' which in turn can be delineated into operational and management processes [Davenport, 2013] enabling the flow of material, information and money through the value chain. Thus, IS research has examined the enablement of execution using IT under research streams such as design [Subramanyam and Krishnan, 2003, Te"eni, 2001, von Alan et al., 2004, Albert et al., 2004, van der Aalst and Kumar, 2003, Purao et al., 2003, Basu and Blanning, 2003, Walls et al., 1992]; development [Krishnan et al., 2000, Harter et al., 2000, Ethiraj et al., 2005, Chiang and Mookerjee, 2004, Morrison and George, 1995, Nunamaker Jr and Chen, 1990, Burstein et al., 1999]; implementation [Baskerville and Wood-Harper, 1998] ; adoption [Venkatesh and Davis, 2000, Szajna, 1996]; usage [Nelson, 1990, Leidner and Kayworth, 2006]; resource based view of IT [Wade and Hulland, 2004]; strategic implications of IT [Sabherwal and Chan, 2001, Piccoli and Ives, 2005]; and business value of IT[Melville et al., 2004].

Third, this model contextualizes the IS discipline by highlighting the boundaries between IS and its adjacent disciplines. Thus, much of the research in scanning and storage technologies typically occur in computer science under research streams of database and sensor network technologies. Research on presentation technologies occurs in schools of information within streams of information visualization and human computer interaction. Research in human interpretation draws heavily from psychology and cognition.

Finally, this framework provides a purposive framework for making sense of the technological advances occurring at once in multiple areas. Industry analysts and practitioners typically view technologies on technological dimensions which themselves are changing at an accelerating rate. We contend that this is akin to finding one's bearing using a compass whose poles are shifting constantly. For instance it is not clear whether the internet and mobile web should be treated as two separate technologies. Thus the interpretive model at once provides an integrative framework while affording purposive categories for the various technological shifts occurring in scanning(sensor networks), storage(noSQL,in-memory computing), presentation(advanced data visualization), interpretation(A.I. technologies), execution(remote actuation) and learning (machine learning, neural nets) stages. As a thought experiment, when nano sensors and actuators become prevalent, we would not need to revamp our framework as they would be easily assimilated in the scan and execute categories respectively. Thus, we contend that the interpretive model provides a stable frame of reference for classifying technological changes.

Having thus developed the interpretive model of IT, we next apply it to tease out the differences between traditional and predictive analytics.

2.5 Traditional and Predictive Analytics Through The Lens of Interpretive IT Model

Seeing Traditional and Predictive Analytics through the lens of the interpretive model of IT reveals clearly the differences between the two (Figure 2.3). While traditional analytics is about presenting data in the form of reports and data visualizations for the human mind to interpret, predictive analytics consists of algorithms that interpret *in situ*. This we contend is the essential difference between the two domains.

In what follows we survey the technological shifts in each of the six stages of scan, store, present, interpret, execute and learn in an effort to tease out the differences between traditional and predictive analytics.

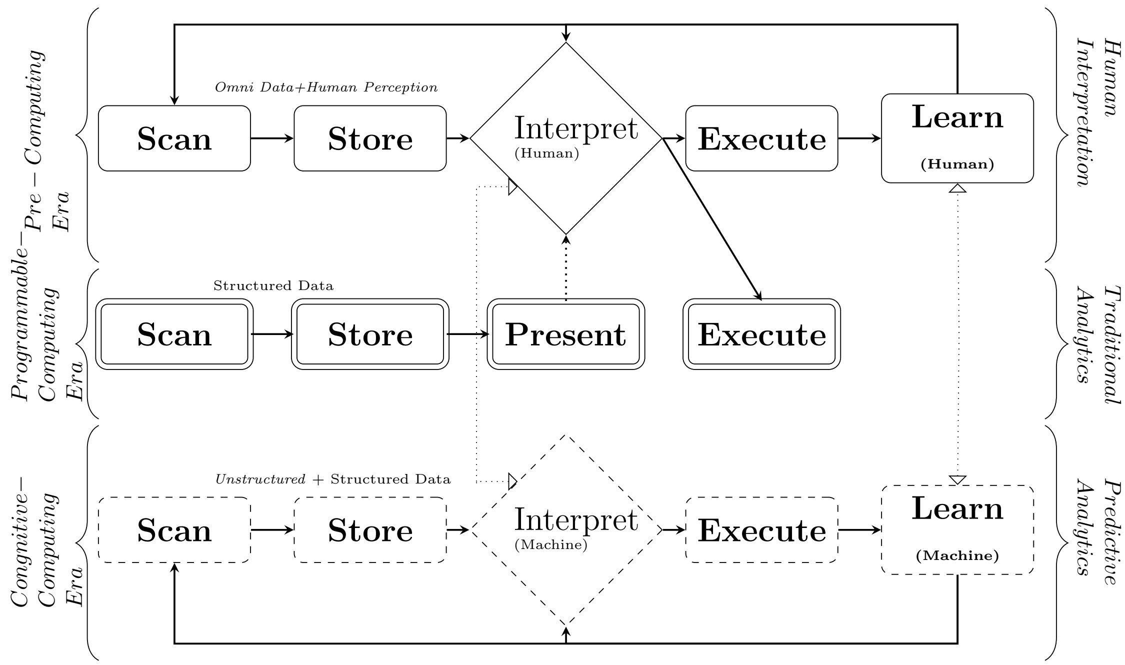


Figure 2.3: Interpretive Model of Information Technology
(Adapted from Daft and Wieck, 1984)

Scan: Scanning of data in the past has been limited to data entered by humans into transaction systems such as ERPs, CRMs and SRMs. However, with a precipitous drop in prices, sensors are becoming ubiquitous. Gartner predicts that by 2020 around 26 billion sensor units will be installed worldwide. These sensors record information about things(internet of things), people(wearable devices,devices embedded inside humans, mobile devices), the terrestrial natural environment(weather sensors, soil sensor networks), the built environment(buildings, bridges, infrastructure) and the extra-terrestrial environment(space telescopes, satellite data). Thus we are moving to a new normal of automated ongoing collection of data streams wherein the percentage of data collected that gets processed is rapidly declining. Though these scanning technologies are ushering in a new era of big data, they are not the primary differentiator between traditional and advanced analytics.

Store: With the ability to scan vast amounts of data about our informational environment, the volume of data to be stored is growing exponentially. Walmart for instance is estimated to collect more than 2.5 petabytes of data every hour from its customer transactions [McAfee et al., 2012]. 90% of all data worldwide was generated in the past two years and 80% of this data is unstructured. New technologies have been developed to meet the challenge of this increase in volume(MapReduce, In-memory databases), variety (NoSQL, Hadoop) and velocity (Databases oriented towards storing time series data such as equipment logs,human health monitoring data, scientific experiment device data). The advances in technology enabling storage of big data have undeniably helped the rise of predictive analytics. However, we contend that storage technologies are an enabler of advanced analytics and not a key differentiator.

Present: We posit that even though traditional business analytics has been called business intelligence, it aids human intelligence rather than possessing any intelligence of its own. Traditional analytics at its core is code fetching and returning records like a dumb robot. Though advances in information visualization help human interpretation of large and complex data sets, what is presented remains bounded by the knowledge that preceded the design of these reports [Tuomi, 1999].

Execute: Technology enables execution by digitizing and automating standardized processes. This is variously termed in the IS literature as process automation [Davenport, 2013] and has been at the core of extant IS research. With the increasing digitization of physical processes, business process automation is being complemented with process automation. Additionally, with the proliferation of sensor and actuator networks, remote actuation such as home automation and smart grids is becoming common place. However, in essence IT enabled execution remains unchanged. What seems new is the role played by information technology in arriving at decisions on what to execute. This crucial shift is discussed next.

Interpret and Learn: As we've been arguing, interpretation together with learning is the distinguishing element of predictive analytics. Traditional analytics is focussed on presentation of information which arguably is the product of knowledge that already existed in someone's mind. In fact, following Tuomi's logic[Tuomi, 1999], human knowledge becomes information once it is articulated and presented in the form of text, graphics or other forms [Alavi and Leidner, 2001]. Predictive analytic systems on the other hand learn from patterns in training data available about a particular scenario. They then start predicting based on the knowledge they have acquired. As they make more predictions and observe the outcomes, they learn and become better at predicting. A common example is the spam filter in our email system. As we tag emails as spam and sometimes recover emails that were wrongly classified, the spam filter in our email becomes better at predicting and identifying spam. Algorithms are typically categorized into predicting continuous quantities (e.g. regression,SVM), predicting clusters(e.g. K-means, Expectation Maximization)and classification(e.g. C4.5,Logistic regression, Naive Bayes).

We summarize the differences between traditional and predictive analytics discussed thus far in Table 2.2.

Table 2.2: Characteristics of Traditional and Predictive Analytics

Characteristics	Traditional Analytics	Predictive Analytics	Human Sensemaking
Scan	Data Entry into Enterprise Systems	Sensors, Mobile devices, Internet, Health devices, Scientific equipment	Human Perception.
Store	RDBMSs	NoSQL, In-memory databases.	Human Memory
Present	Reports, Visualizations	-	
Interpret	-	Algorithms	Human Sense-Making and Interpretation.
Execute	Process Automation	Process Automation and Actuators.	Human Action
Learn	-	Machine Learning Algorithms	Human Learning
Data	Enterprise Transactions	Social Network data, Personal device streams, Web Clickstream, Health data	Perception, Data.
Data Sources	Enterprise Systems, Online Transaction Systems	Internet of Things, Internet of People, Online activity, Scientific experiments	Five Senses , Intuition.

In sum we conclude that though there are quantum shifts in technologies in all six stages, interpretation and learning are the new elements that distinguish predictive analytics from traditional analytics. Further, as the two dotted lines in Figure 2.3 connecting interpretation and learning in humans and machines suggest, we believe in a symbiotic interplay rather than a zero-sum game between man and machine as portrayed in popular and practitioner literature[Brynjolfsson and McAfee, 2014, Brynjolfsson and McAfee, 2011]. These dotted lines also allude to the fact that organizations will vary in how much control humans share with the machine in the interpretation and learning functions. For instance, machine learning can lie anywhere on the continuum from unsupervised, semi-supervised, supervised and reinforced. It is apparent thus that the extent to which firms adopt cognitive technologies such as predictive analytics is an outcome of the interplay between organizational and technological factors. Figure 2.3 is a representation of our understanding of the technological factors at play. In the next section, we weave in organizational factors to complete our theoretical framework and use it to develop our hypotheses.

2.6 Interpretive Model of IT: A Purposive and Stable Taxonomy bringing Order in the Chaos of Technology Changes

A unique feature of the proposed interpretive model of IT is that apart from serving as a theorizing framework, it can serve as a tool for the practitioner to assess a firm's bottlenecks in leveraging advanced analytics. Figure 2.4 points to a possible use of the framework to assess the current state of technology infrastructure and analytics talent - two areas which our study identified as potential bottlenecks.Thus using the framework in Figure 2.4, a firm or an industry can conduct a similar assessment. Such an assessment we believe will be a useful sense-making exercise which might reveal to the firm or industry the areas that are solved problems in its specific case and the areas where it needs to focus its attention and energies. In sum, we believe that the interpretive model apart from being a useful theorizing framework, can also be fruitfully employed by the practitioner.

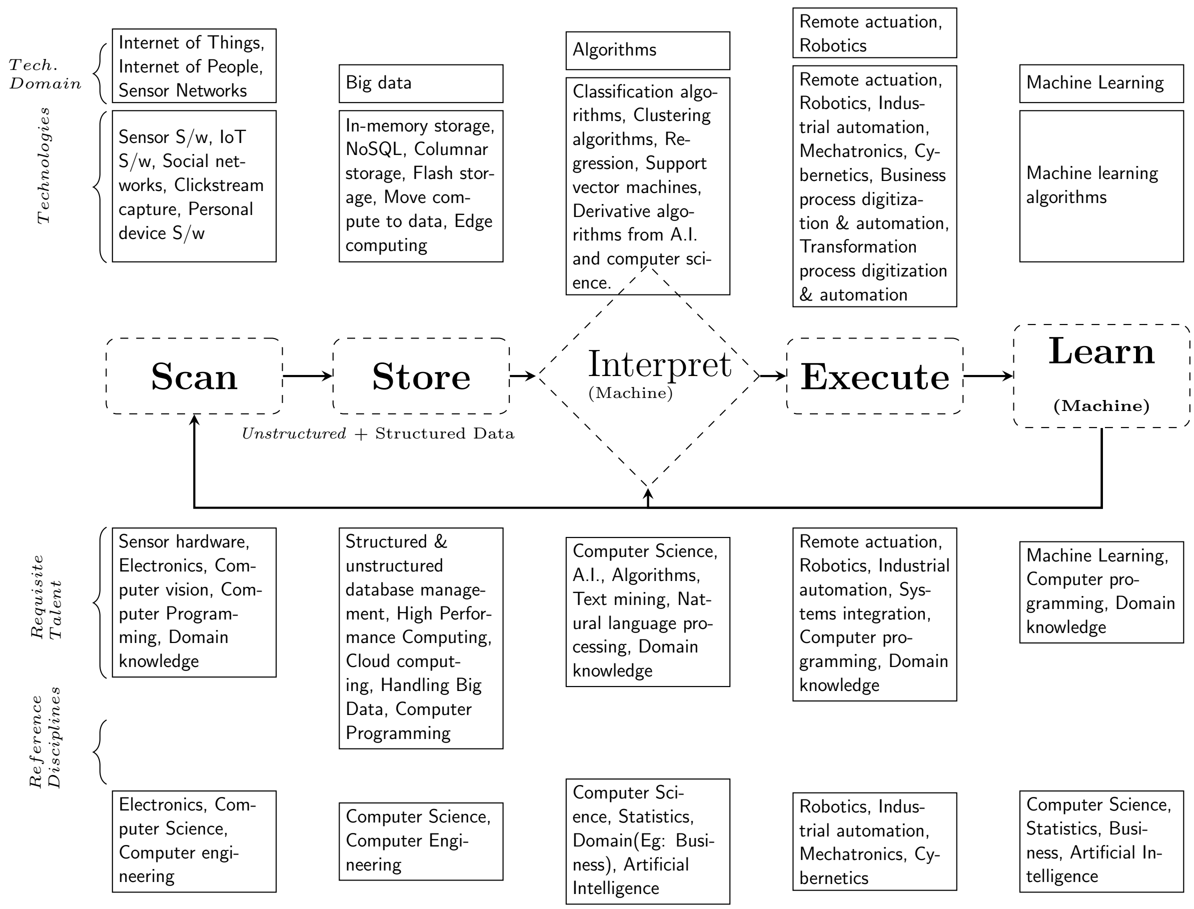


Figure 2.4: Elements of Advanced Analytics parsed through the *Interpretive Model of IT* lens

2.7 Interpretive model of IT: Managerial Implications

The *hype cycle* is a construct of the IT industry analyst community which evaluates the market promotion and perception of value for over 2000 technologies, services and trends categorized under 119 areas. The hype cycle can be considered as an adaptation of the technology adoption curve. In each area technologies are mapped out based on their stage of maturity and the level of expectation from them.

The hype cycle is widely followed by the IT community and is highly influential. This probably owes to the significant benefit it provides to the IT community by collating and providing a snapshot of all technology developments in a single artifact. However, in my opinion its use as a IT decision making aid is questionable. The following introductory quotes from Garner's 2014 special report on hype cycles indicates its position as an IT strategizing lens.

> *The Hype Cycle is a decision aid that will help boards, executive teams, business managers, CIOs,IT leaders, and IT professionals discuss and rationalize the technology and service investment choices in front of them (see "Understanding Gartner's Hype Cycles"). In addition, Hype Cycles can be an invaluable tool to help technology and service providers make effective product planning and marketing decisions (see "Tech Go-to-Market: Using Gartner Hype Cycles to Refine Technology Marketing Strategies and Tactics").*
>
> *What technologies, services and disciplines should you be adopting or*

not? Which options are ready for mainstream adoption, with low risk? What others have yet to be evolved by market feedback and incremental improvement? Are there some things you can use to really lead your industry? If you are a follower, how fast must you move to keep pace? Hype Cycles help organizations determine the appropriate time for them to invest in a technology or service based on their business needs and risk tolerance, rather than based on the market perception and promotion of value. Organizations that use the Hype Cycle to inform strategic investment decisions will see better opportunities in the era of digital business.

It is worth noting that the Hype Cycle focuses on newly emerging technologies as they move into mainstream adoption. Leaving the Hype Cycle does not mean that the technology is being sunsetted; in fact, it may hit a strong period of growth in the mass market. However, the strongest advantage to early adopters and fast-followers is when the technology is still on the Hype Cycle. (For details on the later phases of adoption, including when technologies should be retired, see Gartner's various Market Clocks.

I contend that promoting the hype cycle as an IT investment decision lens is akin to advising a Government to make policy decisions based on the 'now trending' ticker scores on news items. Firms that make their technology investment decisions based on hype cycles are likely to be on the perpetual roller coaster of hypes and troughs of disillusionment. Instead the starting point for a firm's IT investment decisions should be a clear understanding of its business needs. The firm can then derive from its business requirements the capabilities it needs and how IT can provide some of those capabilities.

2.8 Appendix

Appendix 2.A: MISQ and ISR Articles With Subject Containing "intellig*"

Theme⇒		Theorizing				
Year Vol.	Pub.	Title	Authors	Type	IT Artifact	Research Question/Contribution
1999 23.4	MISQ	Explanations from intelligent systems: Theoretical foundations and implications for practice	Gregor, Shirley; Benbasat, Izak.	Theory	KBS	Reviews empirical studies on the nature and use of explanations in knowledge-based systems in light of cognitive effort perspective, cognitive learning theory, and Toulmin's model of augmentation. Conclusions drawn from the review have both practical and theoretical significance.

Theme⇒		Reviews				
Year Vol.	Pub.	Title	Authors	Type	IT Artifact	Research Question/Contribution
1993 4.2	ISR	A classification of information systems: Analysis and interpretation	Ein-Dor, Phillip; Segev, Eli.	Review	Info. Systms.	Seventeen major types of information systems are identified and defined by vectors of their attributes and functions. These systems are then classified by numerical methods.

1995 19.1	MISQ	Early expert systems: Where are they now?	Gill, T Grandon.	Review	Expert Systems	An examination is made of how the first wave of commercial expert systems, built during the early and mid-1980s, fared over time.
2000 11.4	ISR	Research commentary: An agenda for information technology research in heterogeneous and distributed environments	March, Salvatore;et al.	Review	Emerging Tech.	A research agenda is presented for combining computer science, mathematical modeling, systems thinking, management science, cognitive science, and knowledge of organizations and their functions, to meet the challenges of globalization, interactivity, high productivity, and rapid adaptation faced by business organizations.

2007 31.3	MISQ	The Dynamic Structure of Management Support Systems: Theory Development, Research Focus, and Direction.	Clark, Thomas D, Jr; et al.	Theory	MSS	A conceptual model is developed based on an extensive review that encompasses a broad class of systems whose fundamental purpose is the support of managerial actions and decision making.

2007 31.1	MISQ	E-commerce product recommendation agents: Use, Characteristics and Impact	Xiao, Bo; Benbasat, Izak.	Review	R.A.	Based on a review of existing literature on e-commerce RAs, this paper develops a conceptual model with 28 propositions derived from five theoretical perspectives on important aspects of RAs, namely RA use, RA characteristics, provider credibility, and factors related to product, user, and user-RA interaction, which influence users' decision making processes and outcomes, as well as their evaluation of RAs.

1992 16.1	MISQ	Revisiting DSS Implementation Research: A Meta-Analysis of the Literature and Suggestions for Researchers	Alavi, Maryam; Joachim-sthaler, Erich A.	Review	DSS	A quantitative review of the empirical decision support system (DSS) implementation literature is conducted in order to develop guidelines for implementation management and future research.

Theme⇒		Frameworks				
Year Vol.	Pub.	Title	Authors	Type	IT Artifact	Research Question/Contribution
1996 7.3	ISR	The use and effects of knowledge-based system explanations: Theoretical foundations and a framework for empirical evaluation	Dhaliwal, Jasbir S; Ben-basat, Izak.	Frame work	KBS	A model is proposed based on cognitive learning theories to identify the reasons for the provision of KBS explanations from the perspective of facilitating user learning.A 2-part framework is presented to investigate empirically the use of KBS explanations.

1991 15.4	MISQ	An Applied Framework for Classifying the Complexity of Knowledge-Based Systems	Meyer, Marc H; Curley, Kathleen Foley.	Frame work	KBS	Seventeen major types of information systems are identified and defined by vectors of their attributes and functions. These systems are then classified by numerical methods.
1991 15.1	MISQ	Executive Information Systems: A Framework for Development and a Survey of Current Practices	Watson, Hugh J; et al.	Frame Work	EIS	An EIS development framework is presented that includes a structural perspective of the elements of an EIS and their interaction, the development process, and the dialogue between the user and the system.

1991 15.1	MISQ	Executive Information Requirements: Getting It Right	Wetherbe, James C.	Frame work	EIS	An EIS development framework is presented consisting of four techniques - cross-functional systems, joint application design, structured interviewing, and prototyping. These techniques are presented as pragmatic, easy-to-implement solutions for correctly determining executive information requirements.

Theme⇒		Generic Methods & Tools				
Year Vol.	Pub.	Title	Authors	Type	IT Artifact	Research Question/Contribution
2014 38.1	MISQ	Explaining Data-Driven Document Classifications	Martens, David; Provost, Foster	Sys Dev.	Doc. Classfn.	Explain the logic followed by document classification algorithms
2014 25.2	ISR	A Machine Learning Approach to Improving Dynamic Decision Making	Meyer, Georg; et al.	Design	Algo	Develop an approach for improving decision strategies in dynamic environments

2013 37.2	MISQ	Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets During Social Crises	Oh, Onook; et al.	Case Study	Social Media	Using rumor theory, study citizen-driven information processing through Twitter services using data from three social crises: the Mumbai terrorist attacks in 2008, the Toyota recall in 2010, and the Seattle cafe shooting incident in 2012.
2012 36.4	MISQ	A Hidden Markov Model for Collaborative Filtering	Sahoo, Nachiketa; et al.	Sys Dev	Collaborative Filtering	Propose a hidden Markov model to correctly interpret a users' product selection behaviors and make personalized recommendations allowing for a change in user preferences.

2008 32.4	MISQ	The Design Theory Nexus	Pries-Heje, Jan; Baskerville, Richard.	Design	Decision Model	Develop a general method for constructing a design theory nexus as applicable to ill-structured problems characterized by a large degree of uncertainty with respect to how the problem should be approached and how to establish and evaluate the set of alternative solutions.
2000 11.4	ISR	Providing decisional guidance for multi-criteria decision making in groups	Limayem, Moez; DeSanctis, Gerardine.	Empirical	DSS	A feedback and feedforward design is operationalized for a complex multi-criteria modeling system operating within a group decision support system context.
1987 11.4	MISQ	Process Tracing Methods in Decision Support Systems Research: Exploring the Black Box	MISQ	Method	DSS	The applicability to decision support system (DSS) research of process tracing methodologies, particularly verbal protocol analysis, is examined.

1986 10.2	MISQ	Integrating Expert Systems and Decision Support Systems	Turban, Efraim; Watkins, Paul R.	Design	Expert Systems	Possible connections between Expert Systems and Decision Support Systems are discussed. the 2 technologies are examined.

Theme⇒ Specific Applications

Year Vol.	Pub.	Title	Authors	Type	IT Artifact	Research Question/Contribution
2012 36.4	MISQ	Business Intelligence in Blogs: Understanding Consumer Interactions and Communities	Chau, Michael; Xu, Jennifer.	Frame work	Blogs	Propose a framework for gathering business intelligence from blogs by automatically collecting and analyzing blog contents and bloggers' interaction networks
2012 36.4	MISQ	A Social Network-Based Inference Model for Validating Customer Profile Data	Park, Sung-Hyuk; et al.	Sys Dev	Algo	Construct a relational inference model, based on social network and homophily paradigms, to determine the accuracy of self-administered consumer profiles.

2012 36.4	MISQ	Web 2.0 Environmental Scanning and Adaptive Decision Support for Business Mergers and Acquisitions	Lau, Raymond Y K; et al.	Sys Dev	Algo	Design and develop an adaptive business intelligence (BI) 2.0 system underpinned by an evolutionary learning approach, domain-specific sentiment analysis, and business relation mining to operationalize a scorecard model for adaptive M & A decision support.
2012 36.4	MISQ	Network-Based Modeling and Analysis of Systemic Risk in Banking Systems	Hu, Daning; et al.	Sys Dev	Algo	Develop a network approach to risk management (NARM) for modeling and analyzing systemic risk in banking systems viewing banks as a network linked through financial relationships and analyzing systemic risk attributed to each individual bank via simulations based on real-world data from the Federal Deposit Insurance Corporation.

2012 36.4	MISQ	MetaFraud: A Meta-Learning Framework for Detecting Financial Fraud	Abbasi, Ahmed; et al.	Frame work	Fraud Detec- tion	Use a design science approach to develop MetaFraud, a novel meta-learning framework for enhanced financial fraud detection, in an effort to protect firms from financial fraud.
2012 23.3	ISR	From Business Intelligence to Competitive Intelligence: Inferring Competitive Measures Using Augmented Site-Centric Data	Zheng, Zhiqiang (Eric); et al.	Design	C.I. S/w	Develop a method to incorporate competitive intelligence in BI systems by using less granular and aggregate data, which is usually easier to acquire than rich data about a firm's competitors.
2009 20.4	ISR	Designing Intelligent Software Agents for Auctions with Limited Information Feedback	Adomavicius, Gedimi- nas; et al.	Sys Dev	Auction Agents	Present analytical, computational, and empirical analyses of strategies for intelligent bid formulations in online auctions. Strategies are designed for software agents to make bids intelligently.

2006 30.1	MISQ	Incorporating Software Agents into Supply Chains: Experimental Investigation with a Procurement Task	Nissen, Mark E; Sengupta, Kishore.	Expmnt.	Procurement Agents	By investigating the comparative performance of human and software agents across varying levels of ambiguity in the procurement domain, the experimentation described in this article helps to elucidate some new boundaries of computer-based decision making quite broadly.
1997 8.3	ISR	When processes learn: Steps toward crafting an intelligent organization	Zhu, Dan; et al.	Case Study	Intel. Prodn. Schedlr.	The crafting of an organizational process which can learn is presented, and a new set of organizational learning metrics are developed and applied to that process.
1987 11.3	MISQ	A Powerful MIS/DSS Developed for a Remote Sawmill Operation	O'Keefe, J B; Wade, P F.	Case Study	DSS	The development and successful implementation of an integrated, multiapplication system for Juniper Lumber Co. Ltd. of New Brunswick, Canada, is described.

Theme⇒ Adoption & Usage

Year Vol.	Pub.	Title	Authors	Type	IT Artifact	Research Question/Contribution
2013 24.3	ISR	Motivational Differences Across Post-Acceptance Information System Usage Behaviors: An Investigation in the Business Intelligence Systems Context	Li, Xixi; et al.	Empirical	BI Systems	Identify factors affecting post-acceptance information system (IS) usage behaviors
2000 11.4	ISR	Memory-based feedback controls to support groupware coordination	Bordestsky, Alex; Mark, Gloria.	Empirical	Groupware	empirical study is presented of groupware use illustrating problems that users faced with restricted feedback about others' activities.
1995 19.2	MISQ	The impact of explanation facilities on user acceptance of expert systems advice	Ye, L Richard; Johnson, Paul E.	Empirical	Expert System	An empirical investigation into the impact of ES explanations on changes in user beliefs toward ES-generated conclusions.
1992 16.3	MISQ	The Use of Information in Decision Making: An Experimental Investigation of the Impact of Computer-Based Decision Aids	Todd, Peter; Benbasat, Izak.	Expmnt.	DSS	Two experiments were set up to compare the extent of information use by unaided decision makers and users of a decision aid designed to support preferential choice problems.

1990 14.1	MISQ	Expert Systems: A Question Of Liability?	Mykytyn, Kathleen; et al.	Commentray	Expert Systems	In the context of Expert Systems, a framework is presented that traces the development of a product, legal issues related to the development, and normative measures that organizations can take to prevent legal calamities.
1988 12.1	MISQ	How Long Will Computers Stay Dumb?	Kugel, Peter.	Commentray	Intel. Systms.	A contemplation of what technological development might make computers "intelligent".
1986 10.4	MISQ	Toward Intelligent Decision Support Systems: An Artificially Intelligent Statistician	Remus, William E; Kottemann, Jeffrey E.	-	DSS	An expert DSS for statistical analysis, called an artificially intelligent statistician (AIS), is proposed and used in an illustration of unintelligent and intelligent support.

1985 9.2	MISQ	Design and Implementation of Decision Support Systems in the Public Sector	Henderson, John C; Schilling, David A.	Case Study	DSS	An in-depth case study of the design and implementation process of a DSS in a public sector medical facility is presented.

Theme⇒		Business Value				
Year Vol.	Pub.	Title	Authors	Type	IT Artifact	Research Question/Contribution
2013 37.4	MISQ	Differential Influence of Blogs Across Different Stages of Decision Making: The Case of Venture Capitalists	Aggarwal, Rohit; Singh, Harpreet.	Empir ical	Blogs	Empirical investigation of differential influence of online user-generated content (UGC), specifically blogs, across the multiple stages of decision making of venture capitalists: screening stage, choice stage, and contract stage.
2013 37.1	MISQ	Impact of Information Feedback in Continuous Combinatorial Auctions: An Experimental Study of Economic Performance	Adomavicius, Gedimi- nas; et al.	Expmnt.	Real Time DSS for Auction	Compare the economic performance of continuous combinatorial auctions under three progressively advanced levels of feedback from real-time DSS tools.

2006 30.1	MISQ	The Differential Use and Effect of Knowledge-Based System Explanations in Novice and Expert Judgment Decisions	Arnold, Vicky; et al.	Empirical	KBS	This study examines the way users with varying levels of expertise use alternative types of KBS explanations and the impact of that use on decision making.
2000 24.4	MISQ	Is a map more than a picture? The role of SDSS technology, subject characteristics, and problem complexity on map reading and problem solving	Mennecke, Brian E; et al.	Empirical	SDSS	Investigated how the use of a spatial decision support system - a type of geographical information system - influenced the accuracy and efficiency of different types of problem solvers completing problems of varied complexity.

1993 17.3	MISQ	Determining information requirements for an EIS	Watson, Hugh J; Frolick, Mark N.	Empirical	EIS	A multi-stage study was conducted to explore: 1. methods used to determine the information requirements for the initial and ongoing versions of an EIS, 2. how frequently the methods are used, 3. how useful the methods are, and 4. in what situations the methods are useful or not useful.
1990 14.2	MISQ	An Examination of the Impact of Expert Systems on the Firm	MIS Quarterly	Empirical	Expert System	A case study of one of the very few successful applications of expert systems - Digital Equipment Corp. (DEC) use of the XCON expert system to change the management and execution of the knowledge-intensive task of computer configuration.

1990 14.1	MISQ	Comments On 'Price And Value Of Decision Support Systems'	Toraskar, Kranti V; Joglekar, Prafulla N.	Frame work	DSS	A critique of a framework for reconciling the conflicting approaches in cost-benefits literature on decision support systems (DSS)
1988 12.3	MISQ	Using a GDSS to Facilitate Group Consensus: Some Intended and Unintended Consequences	Watson, Richard T; et al.	Expmnt.	GDSS	An experimental study was used to examine the impact of GDSS use on effectiveness of resolving conflicts of personal preference using a task requiring resolution of competing personal preferences in groups.

1988 12.2	MISQ	The Quantification Of Decision Support Benefits Within The Context of Value Analysis	Money, Arthur; et al.	Frame work	DSS	A value analysis is given that proposes a methodology using conjoint measurement for: 1. determining significant intangible benefits for a specified DSS, 2. using value terms to quantify these benefits, and 3. creating a decision rule for determining the significance of the proposed system.
1988 12.2	MISQ	Computer-Based Support For Group Problem-Finding: An Experimental Investigation	Gallupe, R Brent; et al.	Expmnt.	GDSS	A laboratory study was conducted to explore the effects of group decision support systems (GDSS) technology on group decision caliber and individual perceptions within a problem-solving context.

1987 11.4	MISQ	Price and Value of Decision Support Systems	Pieptea, Dan R; Anderson, Evan.	Frame work	DSS	Practical problems in identifying the development cost and potential value of a decision support system (DSS) are examined, and a 2-dimensional framework for cost-benefit analysis is proposed.
1987 11.1	MISQ	The Management Information and Decision Support (MIDS) System at Lockheed-Georgia	Houdeshel, George; Watson, Hugh J.	Case Study	MIDS	A specific MIDS is studied in-depth for its design and benefits.
1985 9.1	MISQ	A Field Study of Organizational Factors Influencing DSS Success	Sanders, G Lawrence; Courtney, James F.	Field Study	DSS	A survey was used to study the contribution of 1. decision context (degree of problem structure), 2. level of task interdependence (degree of interaction with others), and 3. level of task constraints (degree of decision maker authority and autonomy) on the success of DSS as measured by overall satisfaction and decision-making satisfaction.

Theme⇒		Design				
Year Vol.	Pub.	Title	Authors	Type	IT Artifact	Research Question/Contribution
2014 38.2	MISQ	The Nature and Consequences of Trade-Off Transparency in the Context of Recommendation Agents	Xu, Jingjun (David); et al.	Expmnt.	Recommendation Agents	Investigate a novel design for an RA interface that enables it to interactively demonstrate trade-offs among product attribute values
2009 33.2	MISQ	Interactive Decision Aids for Consumer Decision Making in E-commerce: The Influence of Perceived Strategy Restrictiveness	Wang, Wei-quan; Ben-basat, Izak.	Expmnt.	Online decision aids	Experimental comparative analysis of three online decision aids in terms of users' perceptions of strategy restrictiveness, advice quality, and cognitive effort.
2008 10.2	MISQ	Using an Attribute-Based Decision Support System for User-Customized Products Online: An Experimental Investigation	Kamis, Arnold; et al.	Expmnt.	DSS	An experimental study of an alternative-based and an attribute-based DSS for product customization by online customers to examine the mediating role of decision process variables in the use of an online customer DSS.

2003 27.3	MISQ	The influence of query interface design on decision-making performance	Speier, Cheri; Morris, Michael G.	Expmnt.	DSS	Effect of query interface design of decision making tools on decision-making performance is investigated. Specifically, the use of visual and text-based interfaces is compared on both low and high complexity tasks.
1993 17.2	MISQ	Exploring modes of facilitative support for GDSS technology	Dickson, Gary W; et al.	Expmnt.	GDSS	Effectiveness of group decision support systems (GDSS) as a function of two types of GDSS facilitative support are explored: 1. chauffeur-driven, and 2. facilitator-driven.
1993 17.1	MISQ	Cognitive feedback in GDSS: Improving control and convergence	Sengupta, Kishore; Te'eni, Dov	Expmnt.	GDSS	The effect of computer generated cognitive feedback in computer-supported group decision process is investigated.

1991 15.1	MISQ	Decisional Guidance for Computer-Based Decision Support	Silver, Mark S.	Design	DSS	Three aspects of decisional guidance in DSS are examined: 1. when and why system designers should provide decisional guidance, considering the opportunities, motives, and means for guiding, 2. how designers can provide guidance, introducing a 3-dimensional typology for deliberate guidance, and 3. the consequences of decisional guidance.
1989 13.2	MISQ	A Study Of The Relationship Between Decision Model Naturalness and Performance	Kottemann, Jeffrey E; Remus, William E.	Design	DSS	The bootstrapping paradigm of psychological research was used to examine how reliably decision model naturalness and performance are related.

1989 13.1	MISQ	Logic Programming As A Paradigm For Financial Modeling	MISQ	Sys Dev	DSS	Logic programming (PROLOG) is investigated as a vehicle for structuring and implementing decision support systems, with special attention given to those dealing with financial modeling.
1988 12.4	MISQ	A Study Of Influence In Computer-Mediated Group Decision Making	MISQ	Empir ical	GDSS	The effect of influence behavior on quality of decision making in groups is examined and compared between groups using GDSS and those not using it.
1988 12.4	MISQ	Model Management For Group Decision Support	Ting-Peng, Liang	Design	GDSS	The issues in designing and implementing group model management systems (GMMS) are examined. A software architecture is developed that is composed of an inference engine and 3 subsystems.

1987 11.2	MISQ	Behavioral Theories Relating to the Design of Information Systems	Lovata, Linda M.	Theory	Systems	Three theories that predict user reaction to information are reviewed for implications for design of information systems - 1. operant theory, 2. expectancy theory, and 3. garbage can theory.
1987 11.2	MISQ	Semi-Structured Recurring Decisions: An Experimental Study of Decision Making Models and Some Suggestions for DSS	Remus, William; Kottemann, Jeffrey E.	Expmnt.	DSS	An experiment was conducted to analyze the evidence for tracking when making a production scheduling decision. Approaches to support tracking behavior in semi-structured recurring decisions are presented.

1987 11.1	MISQ	Information Intensive Modeling	Orman, Levent.	Design	Systems	The circular relationship between formal information systems and organizational decision models is investigated. Three possible approaches to resolve the problem are presented: 1. normative design, 2. equilibrium analysis, and 3. evolutionary systems, the most appropriate approach.
1986 10.2	MISQ	Decision Support Planning and Analysis: The Problems of Getting Large-Scale DSS Started	Meador, C Lawrence; et al.	Frame work	DSS	A hybrid technique – Decision Support Analysis (DSA) – is presented to building a large-scale DSS supporting multiple users in a large organization.

1986 10.1	MISQ	An Investigation of the Effectiveness of Color and Graphical Information Presentation Under Varying Time Constraints	Benbasat, Izak; Dexter, Albert S.	Expmnt	Systems	A laboratory experiment was undertaken to assess the influence of color and information presentation differences on user perceptions and decision making under varying time constraints.
1985 9.3	MISQ	Information Systems Support for Group Planning and Decision-Making Activities	Rathwell, Margaret A; Burns, Alan.	Design	Systems	The organization of group-planning and decision-making activities is examined for a variety of application areas and it is shown how a distributed decision-making system – a network integrating separate decision-support systems – can be used to effectively coordinate the information and communication needs of planning and decision-making groups.

CHAPTER III

Who Hath the Crystal Ball? Antecedents of Predictive Analytics Usage Within Firms: An Empirical Study

Economics has been described as the science of allocating scarce resources. Allocating scarce resources in organizations can be likened to an ongoing iterative dance between making decisions and executing them. Till recently, Information Technology (IT) has remained limited to execution of tasks with humans staying in control of decision making. However, there seems to be a gradual appropriation by IT of decision making functions heralding the dawn of the cognitive computing era. Some applications of cognitive technologies in the business world have included financial advisory services , customer engagement machine agents, financial crimes prevention and smart grid applications. These applications are powered by prediction of future events such as equipment failure, fraud, sales trends, stock market movements, customer behavior and venture capital deal attractiveness. The competitive advantage

afforded by this sixth sense of sorts is not lost on organizations and explains the uptick in demand for and supply of advanced analytics. It is increasingly evident that the ability to gain forward-looking insight from advanced analytics will differentiate the winners from the losers. However, it is not clear what differentiates firms that are able to leverage these new technologies from those that are overwhelmed or are plain clueless. Specifically, what are the markers of firms that successfully adopt predictive analytics? In this paper we attempt to answer this question by empirically investigating the antecedents of predictive analytics usage within firms. We find that possessing big data is not a sufficient condition for leveraging advanced analytics. Having a clear vision of the business value of big data, having the infrastructure and tools to manage big data and having the right analytics talent seem essential for harnessing this new natural resource. Additionally, there seems to be a path dependence between traditional analytics capability and advanced analytics capability though we have not been able to test it empirically. Managerial implications of our findings are discussed.

Keywords: Interpretive Model of IT, Advanced Analytics, Predictive Analytics, Cognitive Computing, Big Data, Business Intelligence, Interpretive Model of Organizations.

3.1 Introduction

Economics has been described as the science of allocating scarce resources. Allocating scarce resources in organizations can be likened to an ongoing iterative dance between making decisions and executing them. Till recently, Information Technology (IT) has remained limited to execution of tasks with humans staying in control of decision making. However, there seems to be a gradual appropriation by IT of decision making functions heralding the dawn of the cognitive computing era [1]. The world witnessed the cognitive capabilities of machines when Deep Blue defeated the reigning world chess champion Gary Kasparov in 1997, and when Watson singlehandedly defeated two Jeopardy world champions in 2011. Such demonstration technologies are increasingly finding their way into real world applications such as driverless cars, speech recognition, medical diagnostics and computational criminology among others.

Some applications of cognitive technologies in the business world have included financial advisory services [2] , customer engagement machine agents, financial crimes prevention and smart grid applications [3]. These applications are powered by prediction of future events such as equipment failure, fraud, sales trends, stock market movements, inventory demand, customer behavior, fit between person and a job, venture capital deal attractiveness and others. They are variously labeled under

[1] Ginni Rometty, IBM InterConnect, http://youtu.be/iNZj38sD81w (Oct 10, 2013)

[2] You may soon get advice from a machine," CNBC , http://www.cnbc.com/id/101747606, (June 2014).

[3] Martin LaMonica, "Numentas Brain-Inspired Software Adds Smarts to the Grid," http://www.technologyreview.com, (Feb8,2013).

monikers such as advanced analytics, big-data analytics, prescriptive analytics and predictive analytics.

The competitive advantage afforded by this sixth sense of sorts is not lost on organizations and explains the uptick in demand[4] for and supply[5] of advanced analytics. It is increasingly evident that the ability to gain forward-looking insight from advanced analytics will differentiate the winners from the losers[6]. However, it is not clear what differentiates firms that are able to leverage these new technologies from those that are overwhelmed or are plain clueless. Specifically, what are the markers of firms that successfully adopt predictive analytics? In this paper we attempt to answer this question by empirically investigating the antecedents of predictive analytics usage within firms. We find that possessing big data is not a sufficient condition for leveraging advanced analytics. Having a clear vision of the business value of big data, having the infrastructure and tools to manage big data and possessing the right analytics talent seem essential for harnessing this new natural resource. Additionally, there seems to be a path dependence between traditional analytics capability and advanced analytics capability though we have not been able to test it empirically. Managerial implications of our findings are discussed.

[4]Gartner predicts that analytics will reach 50 percent of potential users by 2014 and by 2020, that figure will be 75 percent, http://www.gartner.com/newsroom/id/2510815 (June 8, 2013)

[5]Two-thirds of IBM Research work is now devoted to data, analytics and cognitive computing, Chairmanś letter to Investors, IBM 2013 Annual Report.

[6]Tim McGuire, "Big data and advanced analytics will define the difference between the losers and winners going forward," http://www.mckinsey.com/videos/video vid=2448291043001&plyrid=2399849255001, (March 2013)

In Section 4.2 we start with a review of extant IS literature on business intelligence and analytics which reveals ambiguity about the differences between traditional and advanced analytics in both practitioner and academic literature. In Section 2.4 we propose a new theoretical model - '*Interpretive Model of IT*', to bring out a sharper distinction between the two. Next, in Section 2.5 we apply this new theoretical model to tease out the differences between traditional and advanced analytics. In Section 4.3, we develop a theoretical framework of technology enabled organizational interpretation extending theories of organizational and managerial interpretation and cognition. In Section 3.4 we exploit this framework to develop our hypotheses. In Section 4.6 we discuss our data and variables which we use to specify and estimate our empirical model in Section 4.7. Results are presented in Section **??** and implications discussed in Section 3.7. We discuss contributions and limitations of the study in Section 4.10. Section **??** draws out the conclusions of our study.

3.2 Literature Review

In extant literature, Predictive analytics has been treated as part of traditional business intelligence and analytics [Chaudhuri et al., 2011, Watson and Wixom, 2007]. Practitioner oriented literature too has propagated this monolithic view of Business Intelligence and Analytics(BIA) [Davenport, 2006, LaValle et al., 2013]. Industry analyst reports have further compounded the ambiguity in the field by covering the two technologies under one umbrella[Bitterer, 2011, Sallam et al., 2011] or characterizing the new developments as hype[Taft, 2012]. Such confusion in a field can indeed be signs of a passing fad or that of the punctuated disequilibrium of a field undergoing a paradigm shift [Kuhn, 1962]. We contend that the later is the case. Recently, there have been tell-tale signs of the acceptance of a new field and paradigm [Kuhn, 1962]. Gartner, after covering business intelligence and analytics for seven years as a single technology, in 2013 decided to study and analyze analytics and advanced analytics as separate fields.

Advanced analytics might seem like an extrapolation of traditional business intelligence and analytics. However, we contend that it represents a paradigmatic shift in that it makes the jump from aiding human *insights* to providing *foresights*. Rather than presenting decision makers with reports and visualizations of historic data, predictive analytics expands the boundaries of managerial and organizational cognition by providing humans with fast, efficient access to foresight trapped in huge volumes of unstructured and structured data. Thus the essential difference seems to be between *presentation* and *interpretation* of data. Put in other words, traditional analytics aids while predictive analytics augments human interpretation. We believe that this shift is not trivial and represents an evolutionary shift from IT being a dumb fetcher of records to a smart interpreter of its informational environment. However, the academic IS literature to our knowledge continues to treat BIA as a singular construct. For an extensive review of IS literature on BIA, we direct readers to Chen et. al. [Chen et al., 2012].

Chen et. al. do differentiate between three stages of BIA evolution. This categorization seems primarily based on the nature of data sources and technologies being used in each phase. Their research framework groups BIA research into Big data analytics, Text analytics, Web analytics, Network analytics and Mobile analytics. Their treatment is an excellent synthesis of the historiography and taxonomy of the field. However, we feel that like any other taxonomy, it might be limited in its exhaustiveness and with the discovery of new categories might need to be revised. Further, it is easy to foresee situations where technologies could span more than one category. For instance it is not clear whether internet usage on mobile devices falls under mobile or web analytics. In fact, in a mobile-first world, one might question whether there is a strong enough case to treat these as separate categories.

In sum, we feel that the BIA field is in need of an organizing framework which remains unaffected by constant changes in technologies. In what follows, we make such an attempt by providing a new theoretical framework for IT called the *Interpretive model of IT*. This model rather than being historiographic in nature, is based on a purposive delineation of IT's role in organizational sense-making. We believe that using purposive rather than technological dimensions as the foundations of our framework provide it stability in the face of unabated and continuous technological changes. Additionally, this framework brings into clear focus the differences between traditional and advanced analytics.

3.3 Theory

The Organizational Information Environment: Shifting Boundaries between the Analyzable, Unanalyzable and Cognizable Information Worlds In their interpretive model of the organization, Daft and Wieck identify four different modes that organizations follow to interpret their environment: enacting, discovering, undirected viewing and conditioned viewing. Each of these modes is determined by (1) management's beliefs about the environment and(2) organizational intrusiveness.

Daft and Wieck contend that managers in an organization make assumptions about their environment and categorize them into analyzable and unanalyzable information worlds. It is salient to note here that these are not universal categorizations but depend upon a combination of the characteristics of the environment as well as the organization's previous interpretation experience. Thus it is possible for the same information world to be perceived as analyzable by one firm and unanalyzable by another. The differences in organization's assumptions about the analyzability of their environment is crucial for our theoretical lens as it might explain differences in firm orientations towards the use of advanced analytics to make sense of the environment.

When organizations assume that the external environment is concrete, measurable and determinant, they consider the environment as analyzable. Conversely when the environment is perceived as subjective, difficult to penetrate, or changing (Duncan, 1972), managers will see it as less analyzable (Perrow, 1967; Tung, 1979). In analyzable paradigms *one* correct interpretation is assumed which drives linear thinking and logic leading to a search for clear data, models and solutions. Additionally, there is little equivocality reduction achieved by interpreting the analyzable component of the information world. The lack of equivocality reduction from analyzable information is in line with Tuomi's contention that there is a reverse hierarchy from knowledge to information to data[Tuomi, 1999]. Thus any information parsed through a report is confined to the knowledge that was used to design the report in the first place. Such reports might reduce uncertainty but not equivocality. In the unanalyzable paradigm there is no assumption of *a* correct answer. The interpretation process is less linear, more ad hoc and improvisational.

Since our inquiry is on the interpretive ability of IT, we are interested in the analyzability of the organizational environment when seen through the IT lens. Revisiting our interpretive model of IT (Figure 2.3), it is clear that a part of the organizational environment may be unanalyzable by IT because of reverse salients[Hughes, 1993] in one or more of the stages of scanning(for e.g.: inability to digitize human emotions), storage(for e.g.: inability to store unstructured data digitally), or interpretation(for e.g.: inability to interpret unstructured data). In fact, in D&W's treatment the only information that is analyzable are transactional data from enterprise systems which is in line with the traditional analytics paradigm of structured tables and reports. The implicit assumption in Daft and Wieck's model is that interpretation is the forte of the human mind. They state:

> *..in interpretation, the human mind is engaged.*

While this was true till very recently, technological advances are making hitherto unanalyzable parts of our environment interpretable by IT. Developments in sensor technologies are bridging them close to mimicking all five senses of humans. With increasing digitization and newer sensor technologies capturing more dimensions of our information environment, larger parts of the unanalyzable information world is being replaced by the interpretable information world.

To ascertain whether this machine interpretable information world is commensurable with the analyzable or unanalyzable information worlds we consider their characterization in D&W's work. The analyzable information world implicitly assumes structured data and reports as revealed by their characterization:

> *when the environment is analyzable, a larger percentage of the data will be conveyed through the management information system...; Organization assumes that the external environment is concrete, events and processes are hard, measurable and determinant...; Viewing is conditioned in the sense that it is limited to the routine documents, reports, publications and information systems...; Organization will utilize linear thinking and logic and will seek clear data and solutions.*

On the other hand, the implicit assumptions in the conception of the unanalyzable environment is that they are only knowable through human perception as revealed in some of the following quotes:

> *Managers...relied on information obtained through personal contacts and causal information encounters(Aguilar 1967); ..company gathered information from personal contacts...and informants in other companies...interpretation was based on a variety of subjective cues; Generally, the less analyzable the perceived environment, the greater the tendency for managers to pur-*

sue external information gained from personal contact with other managers.

After extensive consideration, we could not find a home for the machine interpretable information world in either the analyzable or unanalyzable information worlds. We thus suggest an extension of Daft and Wieck's characterization of the information environment by carving out a (machine) interpretable information world distinct from the analyzable and unanalyzable information worlds. This part of the environment can be scanned, stored and interpreted by IT but cannot be analyzed in a predetermined programmatic sense in the vein of RDBMS tables, structured reports and data cubes. Further, the interpretable information world unlike the unanalyzable information world does not require human intervention and is knowable through systems. Finally, these systems are capable of learning similar to humans, except that the *information cycles* of humans are replaced by *machine learning cycles*. We situate this interpretable information world between the analyzable and unanalyzable information worlds of D&W's interpretive model as illustrated in Table 3.1.

Applying the dimension of organizational intrusiveness to the interpretive information world, we posit that some of the firms will be more entrepreneurial and intentional in their use of predictive analytics to seek out opportunities and hence are classified as *Predictors*. Others will be more passive, respond only to crises and are labeled *Reactors*. This completes our theoretical framework as summarized in Table 3.1. In the next section we use this framework as the foundation for developing

Table 3.1: Man-Machine Interpretation Modes (Adapted from Daft and Wieck, 1984)

Relationship Between Interpretation Modes and Organizational Processes		
Unanalyzable	UNDIRECTED VIEWING **Scanning Characteristics:** 1. Data sources: external, personal. 2. Acquisition: no scanning department, irregular contacts and reports, casual information, **Interpretation Process:** 1. Much equivocality reduction 2. Few rules, many cycles **Strategy and Decision Making:** 1.Strategy: reactor. 2.Decision process: coalition building.	ENACTING **Scanning Characteristics:** 1. Data sources: external, personal. 2. Acquisition: no department, irregular reports and feedback from environment, selective information. **Interpretation Process:** 1. Some equivocality reduction 2. Moderate rules and cycles **Strategy and Decision Making:** 1.Strategy: prospector. 2.Decision process: incremental trial and error.
Interpretable (*Machine*) ***ASSUMPTIONS ABOUT ENVIRONMENT***	REACTING **Scanning Characteristics:** 1. Data sources: external+internal, impersonal. 2. Acquisition: Separate departments, Internet of Things, Internet of People, Digital Media. **Interpretation Process:** 1. Much equivocality reduction 2. Cognitive Representation , No Rules, Ongoing self-improving machine learning cycles **Strategy and Decision Making:** 1.Strategy: reactor. 2.Decision process: Algorithmic, A.I., Cognitive technologies.	PREDICTING **Scanning Characteristics:** 1. Data sources: external+internal, impersonal. 2. Acquisition: Separate departments, Internet of Things, Internet of People, Digital Media. **Interpretation Process:** 1. Much equivocality reduction 2. Cognitive Representation , No Rules, Ongoing self-improving machine learning cycles **Strategy and Decision Making:** 1.Strategy: prospector. 2.Decision process: Algorithmic, A.I., Cognitive technologies.
Analyzable	CONDITIONED VIEWING **Scanning Characteristics:** 1. Data sources: internal, impersonal. 2. Acquisition: no department, although regular record keeping and information systems, routine information. **Interpretation Process:** 1. Little equivocality reduction 2. Many rules, few cycles **Strategy and Decision Making:** 1.Strategy: defender. 2.Decision process: programmed, problematic search.	DISCOVERING **Scanning Characteristics:** 1. Data sources: internal, impersonal. 2. Acquisition: Separate departments, irregular special studies and reports, extensive information. **Interpretation Process:** 1. Little equivocality reduction 2. Many rules, moderate cycles **Strategy and Decision Making:** 1.Strategy: analyzer. 2.Decision process: systems analysis, computation.
	Passive	Active

ORGANIZATIONAL INTRUSIVENESS

our hypotheses.

3.4 Hypotheses

Our theoretical framework in Table 3.1 categorizes an organization's information environment into the analyzable, interpretable and unanalyzable. To what extent organizations invest in knowing each of these information worlds depends in part on their perceived need for it. This perceived need is arguably influenced by the organization's strategy. A typology of organizational strategy that speaks directly to the information seeking behavior of organizations are Miles and Snow's Prospectors, Analyzers, Defenders, and Reactors[Miles et al., 1978].Prospectors focus on how to locate and exploit new product and market opportunities and monitor a wide range of environmental conditions and events. Miles and Snow characterize Prospectors as:

> "*Prospector's prime capability is that of finding and exploiting new product and market opportunities; Prospectors scanning activities are not limited to the organization's current domain; Prospector managers typically perceive more environmental change and uncertainty than managers of other two organization types; Prospectors are not suitable to environments where the world of tomorrow is similar to that of today; Prospectors invest heavily in individuals and groups who scan the environment for potential opportunities.*"

Defenders on the other hand seek to create a stable set of customers and products focusing on efficiency. Their information seeking behavior is characterized in Miles and Snow as:

> "*Defenders tend to ignore developments and trends outside of their domains, choosing instead to grow through market penetration; …do little or no scanning of the environment for new areas of opportunity; Defender's primary risk s being unable to respond to a major shift in the market environment; Defenders are ideally suited for an environment where tomorrow's world is similar to today's.*"

Analyzers try to follow a strategy that is a hybrid of Prospectors and Defenders with a surveillance mechanism that is mostly limited to marketing and little R&D. Reactors are firms that follow an inconsistent and unstable strategy which Miles and Snow consider as a failed attempt at following one of the other three strategies.

In our theoretical development we have argued that the analyzable information world is knowable through traditional analytics, the interpretable information world through advanced analytics and the unanalyzable information world through human effort and perception. Further, we have noted that the analyzable world of structured historical data is like a very clear rear view mirror into the past. The interpretable information world on the other hand, though foggy like a windshield, has the potential to provide a glimpse of the future. Thus, the more proactive and entrepreneurial a firm is in seeking out opportunities, the more it will seek out information resident in the interpretable world using advanced analytics. Conversely, the more a firm is focused on sustaining a narrow base of products and customers, the more it will seek out information from the analyzable world using traditional analytics.

Hypothesis 1a. Firms that perceive themselves as prospectors are more likely to use advanced analytics.

Hypothesis 1b. Firms that perceive themselves as defenders are more likely to use traditional analytics.

The Resource based view (RBV) of the firm assumes that firms can be conceptualized as bundles of resources or capabilities which provide them with sustained competitive advantage if the resources are valuable, rare, inimitable and non-substitutable. Resources can be considered as stocks of available factors that are owned or controlled by the firm. *Capabilities*, in contrast, refer to a firm's capacity to deploy the *resources* [Amit and Schoemaker, 1993]. Looking at the analytics domain through the RBV lens, we assert that the resources are data and the technological infrastructure to manage data. Capability on the other hand is the ability to exploit these resources for generating relevant insights and foresight. Requisite capabilities however change with a change in the environment. In proposing the 'dynamic capabilities' approach, Teece et. al.[Teece et al., 1997] stress how firms exploit existing internal and external firm-specific competences to address changing environments.

In analytics, the change in environment is arguably from traditional to advanced analytics requiring a transition in capabilities from managing traditional data through ETL (extract transform load) routines to new capabilities for managing big data.

This transition has been made especially imminent due to the inability of traditional database management tools to handle the volume, velocity and variety of big data.[7].

Researchers of learning theory have found however, that a transition from old to new capabilities cannot be taken for granted. On the contrary, there is evidence for a 'lock-in' in capability building wherein core capabilities are likely to become fixed through a process of self-reinforcement. This binds the organization to the past and excludes the development of newer capabilities.

> "*As organizations develop greater competence in a particular activity, they engage in that activity more, thus further increasing competence and the opportunity cost of exploration*" [Levinthal and March, 1993].

> "*the pitfall is that this learning increases the rigidity of the firm*" [Kogut and Kulatilaka, 2001].

Thus, the capabilities that a firm possesses seem to have a certain level of stickiness, at least in the short term and seem to limit the range of resources a firm can exploit. Thus we contend that only firms that are able to snap out of this stickiness and acquire the requisite capabilities of advanced analytics infrastructure will be able to take advantage of the big data resource and thus adopt advanced analytics.

Hypothesis 2a. Firms that possess big data management infrastructure are more likely to adopt advanced analytics.

[7]Techrepublic, When traditional RDBMS hits the big data performance wall, http://www.techrepublic.com/article/splice-machine-when-traditional-rdbms-hits-the-big-data-performance-wall/ (Aug 19, 2014)

Hypothesis 2b. Firms that possess traditional data management infrastructure are more likely to adopt traditional analytics.

In the domain of big data, the variety of data sources that can be tapped by firms can be enormous. This variety of data sources may range from access to postings on social networks to sensor data streams of every transaction executed in a factory floor and at each step in the supply chain. This variety of data sources increases the opportunity set for the data scientist to derive managerial insights. For example, instead of analyzing trends from simple sales transaction data at retail stores, if a firm has access to store foot-traffic data, social media postings of its customers in that zip code and promotions from competition in that market, the contextual insights that could be generated could be of a qualitatively different nature and significantly more effective. Thus, we contend that possessing big data resource is essential for deriving the foresights that advanced analytics is capable for delivering. Hence we believe that:

Hypothesis 3. Firms that possess Big Data Sources are more likely to use Advanced analytics.

Cohen and Levinthal have defined absorptive capacity (AC) as 'the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends.'(p. 128) [Cohen and Levinthal, 1990]. They further argue that organizational AC develops on prior investments in building AC in individuals; develops cumulatively and tends to be path dependent; and depends on the organizations ability to share knowledge internally. The key role of the AC of individuals in the information technology domain has been widely reported in popular press[8] and recognized in IS literature [Bresnahan et al., 1999, Ang et al., 2011]. IS scholars have extensively researched the impact of IT talent on software quality[Krishnan et al., 2000, Krishnan and Kellner, 1999], software project success[Whitaker et al., 2010], firm capabilities[Ethiraj et al., 2005, Mehra et al., 2014], firm profitability[Bapna et al., 2013], firm's opportunity space[Gopal et al., 2003], customer satisfaction[Ramasubbu et al., 2008], employee compensation[Mithas and Krishnan, 2008] and performance[Bapna et al., 2013].

However, going by the unprecedented nature of the shortage of advanced analytics talent[9], the AC required of data scientists seems unique and rare. This rarity is often attributed to the need to integrate within a single mind the knowledge from three disciplines, each of which require years of training - computer science, statistics and business acumen[Davenport et al., 2013]. A close look at these three domains reveals that computer science and statistics can be considered as *procedural* knowledge while business acumen can be considered as *declarative knowledge*[Anderson, 1996].

This procedural-declarative dual demand on the knowledge corpus of a data scientist is reminiscent of the Adaptive Character of Thought (ACT) theory[Anderson, 1996] , which posits that

> "*complex cognition arises from an interaction of procedural and declarative knowledge.*"

[8]WSJ, CIO Journal,The Dog Fight for Tech Talent, http://blogs.wsj.com/cio/2014/08/18/the-dog-fight-for-tech-talent/ (August 18, 2014)

[9]WSJ, Big Data's Big Problem: Little Talent, http://online.wsj.com/articles/SB10001424052702304723304577365700368073674 (April 29, 2012)

In contrast, enterprise systems industry (Eg: ERPs and CRMs), has achieved a workable division of labor between procedural and declarative knowledge by creating technical(e.g.: software engineers) and functional(e.g.: business analysts) roles respectively. Thus, in a data scientist, the IT industry might be making a first demand to create a role that boundary spans procedural and declarative knowledge across multiple disciplines.

Another unique aspect of the AC of a data scientist is the need to activate the right knowledge components in a given context. This is critical for a data scientist because many of the answers are fuzzy and in many cases the questions themselves are not clear. This demand is usually not placed on IT professionals working on traditional transaction systems, because much of the declarative knowledge has already been codified as procedural knowledge [Anderson, 1982]. This procedural knowledge is usually codified in best practices process templates in enterprise systems, and can be used as-is. A second reason, is that enterprise systems usually operate at the business process level, a level of abstraction that applies across industries. Thus the same accounting, manufacturing and distribution processes can run an oil rig as well as an automotive plant.

Data scientists, on the other hand, need to deeply understand the details and nuances of an industry's value generating processes at layers deeper than that of abstracted business processes. A deeper and wider understanding of the domain in required to derive full benefit from the deeper and wider view available from big data. This assertion is motivated by the *law of requisite variety*, which states that responding effectively to an environment requires a varied enough representation to match the complexity of the environment[Gallagher et al., , Ashby et al., 1956]. Also, ACT theory posits that the power of human cognition depends on the amount of knowledge encoded and the effective deployment of the encoded knowledge[Anderson, 1996]. Thus, as a data scientist, accumulates a deeper and wider knowledge repertoire, her cognitive powers and the richness and diversity of insights she can generate are likely to go up. On the other extreme, uni-dimensional expertise may not be able to derive insights from big data much like a screen with many holes will not display an image projected on it.

Acquiring a deep and wide knowledge base across multiple disciplines is not trivial. It is a labor-intensive enterprise in which one must acquire one-by-one all the knowledge components. Several researchers have validated that acquiring an expert level of competence in a field usually takes a minimum of 10 years of intense practice[Ericsson et al., 1993].Thus, in light of requisite polymathic versatility of a data scientist, the unprecedented shortage of analytic talent[Craig et al., 2012, Davenport et al., 2013] is not surprising.

IS scholars are responding to this call for help through scholarly research[Tambe, 2013, Agarwal et al., 2014], investigating the dimensions of this talent gap[10] and trying to fill it in earnest by starting data science programs. One recent compilation of data science programs [11], has identified 118 colleges and universities offering business

[10] Ryerson University, Canada's Big Data Talent Gap Study, http://www.ryerson.ca/provost/partnerships/ talentgap.html (Oct 20, 2014)

[11] Data science community, Colleges and Universities with Data Science Degrees http://datascience.community/colleges (Oct, 2014)

analytics or data science courses, 81 of which are masters level degrees.The scholarly evidence on the non-triviality of acquiring polymathic competence and the universal lament from industry about analytics talent shortage leads us to the following hypotheses:

Hypothesis 4a. Analytics talent gap is negatively associated with the extent of advanced analytics use in firms.

Hypothesis 4b. Analytics talent gap is negatively associated with the extent of traditional analytics use in firms.

The technology acceptance model(TAM), which is based upon the theory of reasoned action from the social psychology literature [Ajzen and Madden, 1986], postulates that technology adoption behavior is an outcome of two salient beliefs: perceptions of usefulness of the IT and perceptions of ease of use. Empirical testing of this theory has found usefulness to be far more important than ease of use in predicting usage [Davis, 1989, Agarwal and Prasad, 1997]. New information technologies represent innovations for potential adopters [Rogers, 2010]. According to Rogers there are five adopter categories: innovators, early adopters, early majority, late majority, and laggards. In discussing the dominant characteristics of each category, Rogers characterizes innovators as venturesome, early adopters as opinion leaders who are widely respected in their social circle, early majority members as deliberate, the late majority as skeptical about the value of an innovation, and laggards as traditional. In general, early adopters use innovations even when the uncertainty surrounding potential use is high, and the benefits of the innovation have not become widely visible and accepted. Conversely the laggards do not perceive a value in adopting these new technologies and thus hold out till the very end.

Traditional analytics having been at play for more than 20 years is a mature technology with clearly defined value propositions in industry. Thus for most organizations, the value proposition of traditional business intelligence is likely to be clear. Indeed, in the Gartner hype cycle on analytics, traditional business analytics is past its peak. On the other hand, advanced analytics being a relatively new technology, a substantial percentage of firms are likely to suffer from a lack of clarity about its value proposition. Thus we hypothesize:

Hypothesis 5. Lack of clarity about business value of analytics is negatively associated with the extent of advanced analytics use in firms.

3.5 Data and Variable Definition

Our empirical analysis is based on data from the InformationWeek 2014 Business Intelligence, Analytics and Information Management (BIAIM) Survey. Information Week is a leading IT publication and InformationWeek surveys are reliable sources of secondary data used in previous academic studies (For example, Bharadwaj et al. 1999; Whitaker et al. 2007). These surveys target top IT managers in organizations who are in decision-making roles with sufficient overview of their firms IT operations and investments. The 2014 BIAIM survey was conducted online in October 2013 wherein pre-qualified Information Week subscribers were sent an email invitation containing an embedded link to the survey. The respondents were business technology decision-makers at North American companies with significant decision-

making authority and involvement in BIA investments in their organizations. The original dataset comprised of data collected from these decision-makers in 312 firms but only 248 respondents were allowed to complete the survey only if their firm had implemented BIA and if they had significant authority related to BIA purchase and implementation in their organizations. After dropping incomplete or duplicate observations and removing outliers per Cooks distance, (Long and Freese 2003), the final sample comprised of data from 195 firms. The variables are described below.
Dependent Variable(s) **Extent of Predictive Analytics Usage in Business Activities (BIpred) and Extent of Traditional Analytics Usage in Business Activities (BItrdn)**: These two variables are ordered variables indicating the extent of usage of Predictive analytics and Traditional analytics respectively in business activities. Survey respondents were asked "How do you currently utilize analytics or business intelligence? Select all that apply", and were given 14 options - Business activity monitoring, Competitive analysis, Corporate governance, Customer relationship management, Financial analysis, Forecasting, Fraud prevention, Operational process optimization, Predictive analysis,Product development, Product marketing, Risk management, Sales tracking and Social media sentiment analysis.

A close study of these business activities revealed that some business activities were predictive in nature seeking foresight while others were dependent on an 'after the fact' analysis. To ascertain the validity of this categorization, a two factor model was specified in which Fraud prevention, Predictive analysis, Risk management and Social media sentiment analysis loaded onto the latent variable of Predictive Usage of BI (BIpred) while Business activity monitoring, Competitive analysis, Corporate governance, Customer relationship management, Financial analysis and Sales tracking loaded onto the latent variable of Traditional Usage of BI (BItrdn). A complete specification of this two factor confirmatory factor analysis is provided in Appendix 3.10. Both BIpred and BItrdn are summative indices of binaries wherein each element is scored a '1' if Analytics is being used for that respective business activity (1=yes; 0=no). A similar approach was used in Banker et al. (2008).

Independent Variables **Business Needs Orientation - Defenders(Defnd) and Prospectors(Prosp)**: These two factors load onto measures that are captured through two questions related to business needs that were posed to the respondents. The first question asked was - "What factors are driving, or would drive, your organization's interest in using advanced analytics? Please select all that apply", and the responses were: Desire to predict promising new business opportunities (e g , up-sell, cross-sell, best new-customer prospects, promising new products, etc), Desire to identify business risk (customer churn, fraud, default, etc), Desire to optimize business operations (sales, pricing, profitability, efficiency, etc), Need to accurately forecast financial or operating results while there's still time to adjust plans, Need to improve customer understanding and marketing segmentation, Need to stay in compliance with laws or regulatory requirements (money laundering, fair banking, fraud management, etc) and Advanced analytics are not currently a priority for my organization. The second business needs related question was - "What data sources or challenges are driving, or would drive, your organization's interest in doing big data analysis? Please select all that apply", and the responses were: Analyzing high-scale machine data from sensors, web logs, etc, Analyzing social network comments for consumer sentiment, Analyzing Web clickstreams, Finding correlations across multiple, disparate data sources (clickstreams, geospatial, transactions, etc), Identifying computer security risks,Predicting customer behavior, Predicting fraud or financial risk, Predicting product or service sales and Big data analytics are not currently of interest to my organization.

A close analysis of the possible responses across these two questions revealed that organization's perceived business needs from analytics can be classified into those geared towards identifying and exploiting new product and market opportunities as well as predicting threats; and those focused on securing and sustaining existing products and markets. These two correspond with Miles and Snow Prospector and Defender strategy typologies. To ascertain the validity of this categorization, two separate models were specified for both these latent factors. The measures of Optimize Business Operations, Marketing Segmentation and Maintain Compliance with Laws and Regulations were loaded onto the latent variable Defender Orientation of Business Needs(Defnd). The measures of Predict Promising New Business Opportunities , Identify Business Risk , Predict Customer Behavior , Predict Fraud or Financial Risk and Predict Sales were loaded onto the latent variable Prospector Orientation of Business Needs(Prosp). A complete specification of these models is provided in Appendix 3.10.

Extent of Big Data Infrastructure (BDInfra) and Extent of Traditional Data Infrastructure (TDInfra): These two summative factors capture the number and extent of big data management and traditional data management technologies usage within an organization. The respondents were asked “To what extent are the following systems/technologies used within your organization?”, and the options included Cloud-based data mart(s)/warehouses,Cloud-based document/record repository,Complex event processing technology,Data cleansing/data quality tools,Data integration software (ETL),Document imaging/capture (scanning and optical character recognition) , Hadoop, High-scale data mart/data warehouse systems supporting massively parallel processing , NoSQL database(s) , On-premise data mart(s)/data warehouse(s),On-premise document/record repository and Trickle feed/change data capture technologies. For each of the technologies, the possible choices were No current/planned use, Planned use, Used on a limited basis and Used extensively. No current/planned use and Planned use responses were coded as 0. Used on a limited basis responses were coded as ‘1’. Used extensively was coded as ‘2’. This coding approach is informed by past research (Saldanha and Krishnan, 2012). The rationale in defining this variable is that IT infrastructure like data-related infrastructure mirror an organization’s historic progress with the use of IT and tends to be highly path dependent in its accumulation (Keen 1991). As our measure constitutes elements like a firm having systems in place for data warehousing, for master data management and for transforming the data etc., these systems are highly path dependent. Having these systems and capabilities needs prerequisite of specialized capabilities and coordination in terms of infrastructure for data integration and management. Relatedly, firms build sophisticated capabilities for data management before and during the implementation of such initiatives (Wixom and Watson 2001).

A close analysis of these 12 technologies revealed that some of these helped manage the increase volume, velocity and variety of data and thus made up the big data infrastructure of the firm. The rest of the technologies were focused on the traditional data management activities of extract, transform and load (ETL). To ascertain the validity of this categorization, a two factor model was specified in which Complex event processing technology,Hadoop,NoSQL database(s), and Trickle feed/change data capture technologies loaded onto the latent variable of big data infrastructure(BDInfra); while Data cleansing/data quality tools,Data integration software (ETL),Document imaging/capture (scanning and optical character recognition) , High-scale data mart/data warehouse systems and On-premise data mart(s)/data warehouse(s) loaded onto the latent variable of traditional data infrastructure(TDInfra). A complete specification of this two factor confirmatory factory analysis is provided in Appendix 3.10.

Extent of Big Data Sources (BDSources): This factor corresponds to the question "What data sources or challenges are driving, or would drive, your organization's interest in doing big data analysis? Please select all that apply", and the options included 'Analyzing high-scale machine data from sensors, web logs, etc,Analyzing social network comments for consumer sentiment', 'Analyzing Web clickstreams', 'Finding correlations across multiple, disparate data sources (clickstreams, geospatial, transactions, etc)', 'Identifying computer security risks', 'Predicting customer behavior', 'Predicting fraud or financial risk', 'Predicting product or service sales' and 'Big data analytics are not currently of interest to my organization. The first four items correspond to big data sources and are consolidated into a 4-item summative index. Each item was coded as '1' if the organization had a particular data source and '0' otherwise.

Big Data Management Challenges (BDMgtChlng) and Traditional Data Management Challenges (TDMgtChlng): These two latent factors measure the types of data management challenges that an organization faces. The respondents were asked "With your organization's experience in mind, what are your organization's biggest impediments to success related to information management? Please select all that apply". The options included - Accessing relevant, timely, or reliable data, Accessing/managing content such as Word files, email messages and presentations, Cleansing, de-duping or ensuring consistent data,Coping with rapidly increasing volumes of data and/or content, Extracting data/transactional information from paper-based forms and documents, Integrating data (e g , extract, transform, load or data federation), Maintaining reliable and responsive data marts/warehouses, Organizing and maintaining data models and/or taxonomies, Processing high-velocity data streams (e g , financial trade or shipping data),Reducing data latency and supporting faster decision making. Each item within the index was coded as 1 if the organization has faced a particular challenge and 0 otherwise.

A close study of these 10 measures revealed that some of these were challenges related to handling the increasing volume, velocity and variety of data(big data management challenges) while others relate to the traditional challenges of extracting, cleansing , de-duping, transforming, organizing and storing the data (traditional data management challenges). To ascertain the validity of this categorization, a two factor model was specified in which the measures of Accessing relevant, timely, or reliable data,Accessing/managing content such as Word files, email messages and presentations,Coping with rapidly increasing volumes of data and/or content',Processing high-velocity data streams (e g , financial trade or shipping data) and Reducing data latency and supporting faster decision making were loaded onto the latent variable of Big Data Management Challenges (BDMgtChlng) ; while Cleansing, de-duping or ensuring consistent data, Extracting data/transactional information from paper-based forms and documents, Integrating data (e g , extract, transform, load or data federation), Maintaining reliable and responsive data marts/warehouses and Organizing and maintaining data models and/or taxonomies were loaded onto the latent variable of Traditional Data Management Challenges (TDMgtChlng). A complete specification of this two factor confirmatory factory analysis is provided in Appendix 3.10.

Business Analytics Managerial Challenges - Business Analytics Talent Gap (TalentGap); Compatibility Issues (CompIssue); Lack of BA Value Clarity (ValClarityGap): These three factors are related to managerial challenges that organizations face in implementing Business Intelligence and Analytics. These factors are drawn from responses to the question- "What are the biggest barriers to successful analytics or BI initiatives? Please select all that apply". with the following possible responses: Challenges scaling the technology across the entire organization,Challenges getting constituents to agree on standardized product(s), Data quality problems, Ease-of-use challenges with complex software/less technically savvy employees, Integration/compatibility issues with existing/multiple platforms, Lack of industry standards, Lower-than-expected analytic value, No clear ROI,No need for BI capabilities throughout our enterprise, Overlap with other products, Software licenses are too expensive, Talent is too scarce or expensive to hire and Training internal staff too time-intensive and costly.

A close analysis of the possible responses revealed three broad categories of challenges - Lack of talent within the firm to execute business analytics projects (TalentGap), A lack of compatibility with existing standards or technologies (CompIssue), and a lack of clarity within the firm about the benefit of business analytics to the firm (ValClarityGapGap). To ascertain the validity of this categorization, a three factor model was specified in which the measures of 'Talent is too scarce or expensive to hire', 'Training internal staff too time-intensive and costly' and 'Ease-of-use challenges with complex software/less technically savvy employees' were loaded on the latent factor Business Analytics Talent Gap(TalentGap). The measures of 'Challenges getting constituents to agree on standardized product(s)', 'Integration/compatibility issues with existing/multiple platforms', 'Lack of industry standards' and 'Overlap with other products' were loaded on the latent factor Compatibility Issues (CompIssue). The measures of 'Lower-than-expected analytic value', 'No clear ROI' and 'No need for BI capabilities throughout our enterprise' were loaded on the latent factor Lack of BIA Value Clarity(ValClarityGap). A complete specification of this three factor model and confirmatory factor analysis is provided in Appendix 3.10.

Extent of Senior Management Usage of Business Intelligence (SrMgtBIUsg): This 4-item summative factor corresponds to the question "Which of the following users have access to or utilize analytics and BI today? Please select all that apply" and the options included 'C-level corporate executives (including VPs)', 'Customer-service reps', 'Customers', 'Data analysts or data scientists', 'External suppliers/partners', 'Financial managers', 'IT management', 'Knowledge workers', 'Line-of-business managers', 'Sales force', 'All employees' and 'Every employee and partner. We identified four of these 10 user groups as belonging to the senior management category - 'C-level corporate executives (including VPs)','Financial managers', 'IT management' and 'Line-of-business managers'. Each item within the index was coded as '1' if the organization has faced a particular challenge and '0' otherwise.

Control Variables **Organization Size (Size)**: Size in terms of annual revenues. We control for this variable as organizations of different sizes might influence the level of resources that they can muster to implement predictive analytics. Additionally, the need for business analytics within a firm might also be influenced by size related complexity (**cite some study relating organizational size to complexity). Consistent with prior research, we used seven point bracketed variable indicating annual firm revenues (amounts in millions) (1 - less than $6, 2 - $6 $49.9, 3 - $50$99.9, 4 - $100$499.9, 5 - $500$999.9, 6 - $1,000$4,999, 7- $5,000 or more) (Whitaker et al. 2007)

Industry Competitive Intensity (CompIntensity): Competitive intensity of a firm's industry is measured using the four-firm concentration ratio, a commonly used inverse measure for competition (Melville et al. 2007; Porter and Sakakibara 2004). CompIntensity is defined as the sum of the market shares of the top four market share leaders of the firm's industry (Bharadwaj et al. 1999).We use the concentration ratio data provided by the U.S. Census Bureau at the most detailed North American Industry Classification System (NAICS) level for the most recently available year (2007).

Environment Dynamism (EnvDynamism): Informed by past research (e.g., Boyd 1995; Simerly and Li 2000), we operationalized environment dynamism as the standardized variation in industry-level sales revenue over the last 5 years. We regressed annual industry sales data over 5 years for each industry at the 3-digit NAICS industry level against time and divided the standard error of the beta coefficient of the time variable by the average annual sales revenue for each industry to obtain the industry-level index of environmental dynamism

Hi-tech and Telecom industries (HiTechTel): This indicator variable represents whether the firm is in Hi-Tech Industries or Telecom (1=HiTechTel; 0=other). We control for the firms in these two industries as firms in these two industries are at the forefront of BIA adoption and usage (Accenture 2013)

Manufacturing (Manuf): This variable indicates whether the firm's offering is primarily a good or a service (1 = Manufacturing, 0 = Services) (Mithas et al. 2005). This accounts for the possibility that firms in manufacturing or in service industries are more prone to use BIA due to potential differences in the need for agility to meet service needs of customers (Saldanha and Krishnan 2012).

IT orientation (Transformate): Prior research has identified three primary roles for IT in industries automate, informate and transformate, wherein IT is primarily used respectively to automate manual tasks or to provide information for empowering the management or to fundamentally alter ways of doing business (Chatterjee et al. 2001). As done in prior research (Banker et al. 2011), we adopt Chatterjee et al.'s (2001) classification scheme and create a dummy variable that captures transform IT role in the industry. Firms in industries such as airlines, financial services, advertising, information technology, telecom and media etc., were classified as using IT for transformational purposes per Chatterjee et al.'s (2001) and Banker et al.'s (2011) classification. We create this dummy variable to control for firms in such industries where IT is used for transform purposes as these firms are more likely to adopt and use new innovations faster than firms in other industries (Chatterjee et al. 2001) Descriptive Statistics Descriptive statistics and the correlations between the operationalized variables is provided in Tables 3.2 and 3.3 respectively.

Table 3.2: Descriptive Statistics

Variable	Description	Mean	(Std. Dev.)	Min.	Max.	N	VIF
bipred	Usage of Predictive Analytics	0.98	(1.17)	0	4	195	-
bitrdn	Usage of Traditional Analytics	3.28	(2.02)	0	7	195	-
defnd	Defender Orientation of Business Needs	1.83	(1.15)	0	4	195	1.73
prosp	Prospector Orientation of Business Needs	2.06	(1.46)	0	5	195	1.59
tdinfra	Extent of Traditional Data Infrastrucutre	3.59	(2.65)	0	10	195	1.43
bdinfra	Extent of Big Data Infrastructure	0.87	(1.32)	0	7	195	1.24
bdsources	Extent of Big Data Sources	1.21	(1.12)	0	4	195	1.38
tdmgtchlng	Traditional Data Management Challenges	1.9	(1.38)	0	5	195	1.48
bdmgtchlng	Big Data Management Challenges	1.33	(1.2)	0	5	195	1.40
talentgap	Talent Gap	0.95	(0.95)	0	3	195	1.22
compissue	Compatibility Issues with Standards and Technologies	1.07	(1)	0	4	195	1.29
valclaritygap	Lack of Clarity of Value of BI	0.5	(0.7)	0	3	195	1.19
srmgtbiusg	Extent of BI Usage by Senior Management	1.97	(1.35)	0	4	195	1.26
size	Organizational Size	4.44	(1.89)	1	7	178	1.28
envdynamism	Environmental Dynamism Within Industry	29.66	(38.59)	2.53	188.61	195	1.17
hitechtel	Whether Hi-Tech or Telecom Industry	0.2	(0.4)	0	1	195	1.16
transformate	Transformational Role of IT	0.26	(0.44)	0	1	195	1.29
manuf	Whether Manufacturing Industry	0.24	(0.43)	0	1	195	1.30
compintensity	Competition Intensity Within Industry	23.72	(19.14)	1.9	94.40	195	1.18
						Mean VIF	1.33

Table 3.3: Correlation Matrix

Variables	bipred	bitrdn	talentgap	compissue	valclaritygap	srmgtbiusg	tdinfra	bdinfra	bdsources	tdmgtchlng	bdmgtchlng	defnd	prosp	Manuf	compintensity	envdynamism	HiTechTel	transformate	size
bipred	1.00																		
bitrdn	0.55	1.00																	
	(0.00)																		
talentgap	-0.14	-0.14	1.00																
	(0.04)	(0.06)																	
compissue	-0.03	0.02	0.12	1.00															
	(0.64)	(0.80)	(0.09)																
valclaritygap	-0.22	-0.15	0.17	0.23	1.00														
	(0.00)	(0.04)	(0.02)	(0.00)															
srmgtbiusg	-0.00	0.12	0.16	0.20	-0.01	1.00													
	(1.00)	(0.09)	(0.03)	(0.01)	(0.85)														
tdinfra	0.16	0.32	0.15	0.05	-0.02	0.27	1.00												
	(0.02)	(0.00)	(0.03)	(0.47)	(0.80)	(0.00)													
bdinfra	0.23	0.26	-0.05	0.08	0.01	0.06	0.32	1.00											
	(0.00)	(0.00)	(0.49)	(0.26)	(0.88)	(0.42)	(0.00)												
bdsources	0.09	0.13	0.19	0.23	0.01	0.33	0.21	0.15	1.00										
	(0.23)	(0.07)	(0.01)	(0.00)	(0.92)	(0.00)	(0.00)	(0.03)											
tdmgtchlng	-0.05	0.02	0.24	0.31	0.12	0.31	0.21	0.06	0.29	1.00									
	(0.49)	(0.79)	(0.00)	(0.00)	(0.10)	(0.00)	(0.00)	(0.41)	(0.00)										
bdmgtchlng	-0.01	0.03	0.11	0.31	0.16	0.20	0.08	0.15	0.27	0.35	1.00								
	(0.89)	(0.64)	(0.13)	(0.00)	(0.03)	(0.01)	(0.29)	(0.03)	(0.00)	(0.00)									
defnd	-0.01	0.17	0.24	0.22	0.17	0.30	0.27	0.03	0.28	0.43	0.31	1.00							
	(0.89)	(0.02)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.72)	(0.00)	(0.00)	(0.00)								
prosp	0.16	0.35	0.25	0.19	0.09	0.27	0.21	0.04	0.38	0.33	0.29	0.55	1.00						
	(0.02)	(0.00)	(0.00)	(0.01)	(0.22)	(0.00)	(0.00)	(0.54)	(0.00)	(0.00)	(0.00)	(0.00)							
Manuf	-0.00	0.12	0.00	-0.04	-0.12	0.03	-0.06	-0.01	-0.03	-0.10	-0.04	0.01	-0.02	1.00					
	(0.99)	(0.11)	(0.98)	(0.58)	(0.10)	(0.69)	(0.43)	(0.91)	(0.71)	(0.17)	(0.54)	(0.91)	(0.76)						
compintensity	0.09	0.18	-0.05	0.01	-0.16	0.11	0.01	0.04	-0.02	-0.00	0.01	0.11	0.14	0.27	1.00				
	(0.23)	(0.01)	(0.45)	(0.85)	(0.03)	(0.14)	(0.90)	(0.62)	(0.76)	(0.97)	(0.86)	(0.13)	(0.06)	(0.00)					
envdynamism	0.03	-0.02	-0.05	-0.11	0.02	-0.03	-0.06	-0.07	-0.08	0.02	0.08	-0.02	-0.04	-0.14	-0.16	1.00			
	(0.70)	(0.78)	(0.51)	(0.13)	(0.77)	(0.66)	(0.41)	(0.31)	(0.29)	(0.75)	(0.25)	(0.81)	(0.56)	(0.05)	(0.02)				
HiTechTel	-0.04	0.17	-0.06	0.13	0.08	-0.06	-0.01	0.17	-0.05	-0.02	0.09	-0.05	0.07	-0.04	0.02	-0.17	1.00		
	(0.62)	(0.02)	(0.43)	(0.07)	(0.24)	(0.43)	(0.94)	(0.02)	(0.52)	(0.80)	(0.23)	(0.50)	(0.34)	(0.62)	(0.77)	(0.02)			
transformate	0.08	0.08	-0.06	0.12	-0.03	0.01	0.10	0.01	-0.02	0.15	0.06	0.07	0.11	-0.33	-0.10	-0.06	-0.06	1.00	
	(0.26)	(0.29)	(0.42)	(0.09)	(0.66)	(0.88)	(0.15)	(0.93)	(0.74)	(0.04)	(0.39)	(0.36)	(0.14)	(0.00)	(0.15)	(0.41)	(0.41)		
size	0.01	0.07	-0.14	-0.05	-0.05	0.03	0.27	0.06	0.00	0.07	-0.15	0.10	0.05	0.15	0.14	-0.08	-0.08	0.12	1.00
	(0.91)	(0.36)	(0.06)	(0.53)	(0.48)	(0.74)	(0.00)	(0.44)	(0.96)	(0.33)	(0.05)	(0.16)	(0.54)	(0.05)	(0.06)	(0.32)	(0.30)	(0.10)	

Values in parentheses represent p-values

3.6 Empirical Model and Estimation

We consider two dependent variables - degree of predictive analytics usage(BIpred) within firms and degree of traditional analytics usage(BItrdn) within firms. For each of these dependent variables, we develop a cross-sectional model to test our hypothesis. In the first model, the dependent variable is a summative index signifying the degree of usage of predictive analytics within a firm's business activities. In the second model, the dependent variable is a summative index signifying the degree of usage of traditional analytics within a firm's business activities. The differentiation between these two types of business analytics usage has been teased out in detail in our theory development as well as dealt extensively in section 4.6 and Appendix 3.10.We estimate our empirical model using two alternative approaches. In the first approach we consider the two dependent variables as ordered and use ordered logit to regress them independently. In the second approach, we consider the two dependent variables as continuous and regress them jointly using seemingly unrelated regression (SUR) method.

Ordered Logit Method: In our first approach, both the dependent variables are considered as ordered variables. It may be argued that these variables are count variables. But count variables indicate how many times something of similar nature has happened (Long and Freese 2003). For example, these models are used to study number of patents and number of products etc. and each patent or product is considered to have an equal impact weight in additive count variable. In this study, we study the degree of usage of predictive and traditional analytics in organizational functions. Hence for each firm, BIpred consists of 13 levels based on adoption and can take any value between zero and twelve based on usage. The categories in this variable can be ranked, but the distances between the categories are unknown. Hence the weight of each item in the index may not be same as in count variables (Greene 2008). Hence we treat the dependent variable as ordered. A similar measurement approach was used in Banker et al. (2008) and Bardhan et al. (2007). Since the dependent variable is ordered, we use ordered logistic regression for estimation. Ordered Logistic or Ordered Probit models are used in estimation when the dependent variable is ordered (Greene 2008). We control for industries using IT for transformation purposes as the firms in these industries adopt new technologies early towards strategic benefits (Banker et al. 2011). We control for firms in Hi-Tech and Telecom industries at the 3-digit NAICS level as these industries are at the forefront of BIA adoption (Accenture 2013). We also control for firms in manufacturing industries (Mithas et al. 2005).

Seemingly Unrelated Regression Method: Both traditional and advanced analytics though paradigmatically different technologies require common competencies such as database management. Theory of Absorptive capacity suggests that organizational AC develops cumulatively and tends to be path dependent. Further underlying common factors might drive the adoption of traditional and advanced analytics. We believe therefore that there is a possibility of contemporaneous error terms between the above two models. Following this line of reasoning, we estimate the above two full models jointly using seemingly unrelated regression. Results of Ordered Logit and SUR regressions are presented in the next section.

3.7 Results and Managerial Implications

Results of both the Ordered Logit regression and the Seemingly unrelated regression are provided in Table 3.4. Results of both the methods are similar in the direction of the coefficients as well as the significance levels. Further our intuition that the error terms for adoption of traditional and advanced analytics might be correlated is validated by the Breusch-Pagan test of independence. For discussion purposes, we consider the results of SUR estimation and have summarized the significant variables from this model in Table 3.5 for ease of reference. A summary of the results from our hypotheses testing is presented in Table 3.6. In what follows, we present an analysis of the results, and identify new research questions that this analysis brings forth.

Business needs orientation: Defenders and Prospectors The coefficient of *prosp*(a *prospector* orientation of business) is positively significant, thus supporting our hypothesis *H1a*. We hypothesized that firms with a prospector orientation are more likely to use advanced analytics and it is supported. Interestingly, these firms are also more likely to use traditional analytics. This propensity of prospectors to adopt analytics might be due to two potential reasons. Firstly, a prospector orientation in firms might be associated with underlying drivers that cause a firm to be enthused about using analytics to augment business strategies and operations. Thus the same prospectors that are likely to adopt advanced analytics today were possibly the early adopters of traditional analytics in the past. Earlier, they were making do with traditional analytics in an effort to predict the future in the proverbial sense of *'driving ahead while looking in the rear-view mirror'*. This tendency to predict the future

Table 3.4: Regression Results

	Ordered Logit			Seemingly Unrelated	
Variable	BIpred	BItrdn		BIpred	BItrdn
defnd	-0.198	0.032		-0.106	0.004
	(0.167)	(0.15)		(0.089)	(0.139)
prosp	0.315**	0.496****		0.192***	0.472****
	(0.126)	(0.123)		(0.069)	(0.107)
bdinfra	0.257**	0.306***		0.149**	0.257**
	(0.122)	(0.119)		(0.068)	(0.106)
tdinfra	0.093	0.17***		0.043	0.149***
	(0.064)	(0.063)		(0.036)	(0.055)
tdmgtchlng	-0.071	-0.095		-0.067	-0.067
	(0.13)	(0.121)		(0.069)	(0.108)
bdmgtchlng	-0.06	-0.136		-0.043	-0.123
	(0.139)	(0.129)		(0.078)	(0.121)
bdsources	0.146	0.067		0.082	0.087
	(0.152)	(0.143)		(0.084)	(0.131)
srmgtbiusg	-0.08	0.036		-0.048	0.028
	(0.119)	(0.11)		(0.065)	(0.102)
talentgap	-0.292*	-0.474***		-0.198**	-0.458****
	(0.164)	(0.154)		(0.089)	(0.138)
compissue	-0.005	-0.076		0.007	-0.074
	(0.161)	(0.151)		(0.088)	(0.138)
valclaritygap	-0.376*	-0.251		-0.244**	-0.224
	(0.225)	(0.2)		(0.120)	(0.188)
envdynamism	0.004	0.004		0.002	0.004
	(0.004)	(0.004)		(0.002)	(0.004)
compintensity	0.011	0.007		0.004	0.006
	(0.008)	(0.008)		(0.005)	(0.007)
hitechtel	-0.222	0.8**		-0.209	0.733**
	(0.367)	(0.352)		(0.207)	(0.321)
manuf	-0.221	0.509		-0.029	0.519
	(0.382)	(0.347)		(0.208)	(0.322)
size	-0.066	-0.092		-0.044	-0.076
	(0.086)	(0.082)		(0.047)	(0.074)
transformate	0.204	0.295		0.124	0.288
	(0.355)	(0.347)		(0.202)	(0.313)
Log Likelihood	-220.95	-327.83	Log Likl.	-578.01	
LR χ^2	31.44	64.27	χ^2	36.01	78.86
Psuedo-R^2	0.07	0.0893	R^2	0.1683	0.3070
Observations	178	178		178	178
Prob > Chi-square	0.02	0			
			Breusch-Pagan	$\rho = 0.47$, Chi2 = 39.28, p = 0.0	

**** p ≤ 0.001 *** p ≤ 0.01 ** p ≤ 0.05 * p ≤ 0.1

Bivariate ordered probit model yields results consistent with ordered logit and SUR models

Table 3.5: Regression Results Summary

Variable	Description	Predictive Analytics Usage	Traditional Analytics Usage	Hypotheses
prosp	Prospector Orientation of Business Needs	(+)***	(+)****	H1a
defnd	Defender orientation of Business Needs			H1b
tdinfra	Traditional Data Infrastructure		(+)**	H2b
bdinfra	Big Data Infrastructure	(+)**	(+)**	H2a
bdsources	Degree of big data sources			H3
talentgap	Lack of talent to manage analytics	(-)**	(-)****	H4a,b
valclaritygap	Lack of Clarity on Value of Analytics	(-)**		H5
hitechtel	High Tech or Telecom Firm		(+)**	

Table 3.6: Summary of Hypotheses

No.	Hypothesis	Supported?
1a	Firms that perceive themselves as prospectors are more likely to use advanced analytics.	Yes
1b	Firms that perceive themselves as defenders are more likely to use traditional analytics.	No
2a	Firms that possess big data management infrastructure are more likely to use advanced analytics.	Yes
2b	Firms that possess traditional data management infrastructure are more likely to use traditional analytics.	Yes
3	Firms that possess big data sources are more likely to use advanced analytics.	No
4a	Analytics talent gap is negatively associated with the extent of advanced analytics use in firms.	Yes
4b	Analytics talent gap is negatively associated with the extent of traditional analytics use in firms.	Yes
5	Lack of clarity about business value of analytics is negatively associated with the extent of advanced analytics use in firms.	Yes

based on past data might explain high profile guffaws such as the $2.25 billion inventory write-off by Cisco in 2001 [12]. A smaller but more exoteric case makes apparent the widespread destruction of value due to wrong decisions based on historical data. In 2013, one of Australia's largest wine companies,Treasury Wines, poured down $35 Mn. worth of aged and unwanted wine stock down the drain and handed over another $40 Mn. in discounts to middlemen[13]. With the ushering in of advanced analytics it is possible to harness the newly found predictive power of advanced analytics. For instance, one of United States largest farmer cooperatives, Southern States Cooperative, was able to use predictive analytics to proactively identify slow

[12] CNet.com, Cisco's $2.25 bn. mea culpa,http://news.cnet.com/2100-1033-257278.html (May 9, 2001)

[13] The Sunday Morning Herald, Treasury Wines chief defends $160m write-off, http://www.smh.com.au/business/retail/treasury-wines-chief-defends-160m-writeoff-20130715-2pz23.html, (July 15,2013)

moving inventory to maintain sales level with 31% less inventory. Prospectors were always interested in predicting the future, now they have the crystal ball.

A second possible reason for the prospectors undertaking both traditional and advanced analytics might be a path dependence in the capabilities needed for traditional and advanced analytics. Indeed, IS research has found evidence for a path dependence in dynamic information technology capabilities [Lim et al., 2011]. Practitioners too explicitly acknowledge the influence of path dependence in adoption of technologies. IBM for instance considers technological path-dependence a critical factor influencing adoption of cross-channel integration by retailers[14]. However, because our data is cross-sectional, we do not have a means to empirically test the path dependence between traditional and advanced analytics. This is planned as a future extension of our current study.

The coefficient of *defnd*(a *defender* orientation of business) is not significant, thus not supporting our hypothesis *H1b*. Our hypothesis that firms with a defender orientation are likely to be associated with use of traditional analytics was based on the premise that such firms are likely to use historical data in order to better defend their existing turf. The lack of support of this hypothesis might be because in defending their current turf, these firms are fixated on efficient execution of their current strategy and processes to the extent that they do not feel the need to question whether they are executing the right strategy. Such firms are likely to be focused on improving operational efficiency and driving down costs by efficiently executing standardized

[14]IBM, Rate of adoption of cross-channel integration by retailers driven by technological path-dependence, https://www-03.ibm.com/products/retail/services/services.html (Dec., 2014)

processes. In this, they might find use for transactional systems, such as ERPs, and may not feel the need for any analytics functionality - traditional or advanced. Such firms have stayed away from traditional analytics and the same mindset is likely to keep them away from advanced analytics.

A potential barrier to analytics adoption by defender firms is revealed in a recent survey of midsize, c-level executives conducted by *Inc.* for IBM. The survey[15] finds that the need to integrate analytics software with existing enterprise systems such as ERPs and CRMs is considered as an unreasonably high barrier to cross. However, as analytics vendors increasingly reduce the barriers to adoption by improving ease of use (by enabling app like drag and drop functionality) and reducing cost (by provisioning them on the tap over the cloud), it remains to be seen whether defenders will also start leveraging the benefits of analytics. Additionally, as constant change becomes "the new normal" as popularized by works such as the innovator's dilemma [Christensen, 1997], defending turf may no longer be congruent with efficient execution of standardized processes. By the same reasoning it is likely that defenders belonging to industries with higher clock-speeds of change might be the first to bite the analytics bullet.

Having the right tools: Traditional Data Infrastructure and Big Data Infrastructure The coefficients of *bdinfra*(extent of big data infrastructure) and *tdinfra*(extent of traditional data infrastructure) are positively significant, thus supporting our hypothesis *H2a* and *H2b* respectively. We hypothesized that firms possessing traditional data infrastructure are more likely to use traditional analytics and those possessing

[15] Inc. Research(exclusively for IBM), Mid-Market Cloud Computing and Business Analytics Survey, (October 2012)

big data infrastructure are more likely to use advanced analytics. A support for this hypothesis is consistent with our RBV theoretical foundations. Additionally, in a similar vein as the results in the previous section, firms that possess big data infrastructure are also more likely to use traditional analytics. However, firms that possess traditional data infrastructure do not necessarily adopt advanced analytics. Thus, this seems to suggest that possessing traditional data infrastructure is a necessary but not sufficient condition for graduating to advanced analytics.

These findings are reminiscent of *dynamic capabilities* which has been defined as:

> "*...the firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments. Dynamic capabilities thus reflect an organization's ability to achieve new and innovative forms of competitive advantage...*"

It seems thus that there are complementaries between traditional and advanced analytic capabilities as discussed in the previous section. However, finding evidence for these complementarities, may not be sufficiently useful to the practitioner. In academic literature, theoretical, historical and empirical studies of a form of complementarity - path dependence, run the gamut from the choice of technologies [Arthur, 1994, Page, 1999], to the formation of languages and law [Hathaway, 2001]. Surely the micro-level processes that produce path-dependent government policies differ from those that determine the acquisition of technological capabilities. Thus we propose that to be truly useful to the firm aspiring to use advanced analytics to amplify its value creation potential, the IS community needs to take on the more ambitious research goal of investigating the micro-level mechanisms that drive these complementarities between traditional and advanced analytics.

Absorptive capacity: Analytics Talent Gap The coefficient of *talentgap*(analytics talent gap) is significantly negative, thus supporting our hypotheses *H4a* and *H4b*. We hypothesized that a lack of analytics talent is negatively associated with both advanced analytics usage and traditional analytics usage. A support for both these hypotheses seems to echo predictions of the impending shortage of data scientists[Davenport and Patil, 2012]. As we discussed in the previous section, data by itself is not a sufficient condition for leveraging advanced analytics. The key competitive advantage in this scenario seems to be the ability to do the heavy lifting of handling this data and making sense of it. This ability is acquired by bringing together the knowledge of statistics, computer science and business acumen - an infamously rare combination. In that sense, Data scientist seems to be the quintessential valuable, rare, inimitable resource of the moment.

Both academia and industry have responded to this lack of data scientists by opening up data science programs(119 in the US alone by one count[16]) and data science fellowships respectively. These programs are attempting to train students to work as data scientists in industry. However, it is too early to make the call on to what extent they will succeed in developing the right skill set required by industry. A characteristic of a good data scientist as identified by IBM is:

> *Good data scientists will not just address business problems, they will pick the right problems that have the most value to the organization*[17]

[16]Data science community, Colleges and Universities with Data Science Degrees http://datascience.community/colleges (Oct, 2014)

[17]IBM, What is a data scientist, http://www-01.ibm.com/software/data/infosphere/data-scientist/ (Dec, 2014).

Thus apart from the science, there seems to be an element of artistry involved in clearly seeing and extracting the value from big data. Thus a potential research question that needs to be answered is whether the university programs are producing the talent that the industry needs. A Second open question is evaluating which is better preparation for industry - university programs or industry boot camps which are much shorter but more hands on.

As firms grapple with the talent shortage within, they are increasingly looking outside their four walls to fill this talent gap. The open data movement started out with Governments and foundations placing large swaths of their data in the public domain in the spirit of transparency. With the coming of age of data science, this movement has morphed into data science competitions wherein organizations open a relevant part of their data in the public domain and teams compete to improve an objective function using that data. The poster child for this open data science model is the Netflix prize[18] which gave \$1Mn. to the team that substantially improved the accuracy of prediction of movie preferences of users. To formalize such competitions, data science competition platforms such as Kaggle have mushroomed to enable harnessing the wisdom of the crowds. Nasa, GE, Facebook, Careerbuilder and a retailer with \$10B+ revenue are just some of the firms that have successfully leveraged the wisdom of the data science crowd. In light of this 'open data science movement', an open question is the determination of the limits to which a firm can outsource its data science talent to the crowd. Which skills must necessarily remain within the firm? A different question reminiscent of increasing returns and lock-in

[18] http://www.netflixprize.com/

by dominants is whether this shift towards open data science democratizes access to foresights for smaller firms, or further entrenches the position of the dominants by making available to them another tool to concentrate power in their favor - the power of foresight.

Clarity of business value comes first: Value Clarity Gap The coefficient of *valclaritygap*(lack of clarity of value of business intelligence and analytics) is significantly negative, thus supporting our hypothesis *H5*. We hypothesized that a lack of clarity on value of advanced analytics is negatively correlated with its usage is supported. Indeed evocations of despair from industry decrying having to drink from the fire-hose of big data seem to suggest a shift from the constraints of data poverty to the disorientating side-effects of a data deluge. This might also explain the significant scoffing at the 'big data exhaust problem' from many quarters. Paradoxically, too much data seems to be blurring the field of vision for the untrained eye. An interesting question to address is the markers of firms that are able to maintain their bearings by virtue of their clarity of vision. Further, it will be intriguing to explore whether this clarity of vision or the lack of it is rooted in understanding technological dimensions or in understanding the firm's mission, vision, revenue model, business model and operational model.

Data as the new natural resource? Big Data Sources The coefficient of *bdsources*(extent of big data sources) is not significant, thus not supporting our hypothesis *H3*. We hypothesized that possessing big data is positively associated with the use of advanced analytics. However, even through there is sufficient variation in how much big data firms possess, it does not seem to account for any variation in the adoption of advanced analytics. The lack of support for this hypothesis seems surprising with the popular and practitioner media hailing big data as the new natural resource powering the economy [19] implicitly through the use of advanced analytics. However, as we contemplate on the increasing pervasiveness of big data, in a paradoxical sense, its omnipresence is probably rendering it as a *sine qua non* for advanced analytics, much as air is for life. Thus, as we tend towards ubiquitous sensor networks and a shift towards open data the percentage of data collected that can be processed by firms will steadily decrease. In sum, we surmise that data by itself is unlikely to be a competitive advantage in the usage of advanced analytics. Though everyone will probably have access to it, not all firms will be able to make use of it. A research question to address is what capabilities are critical for firms to make good on the big data promise. In the previous sections we have seem empirical evidence for some candidate enabling capabilities - big data and analytics infrastructure, analytics talent and clarity on the business value of analytics. Thus, it might be fair to say that an excessive focus on acquiring big data might serve as a red herring. Firms should first spend their cognitive energies on clarifying the business value of analytics in the specific context of their vision and strategy. Armed with this clarity of value of analytics, they should acquire the requisite infrastructure and talent among other capabilities that will enable them to harness the potential of big data. Indeed, disillusionment with big data, expressed in phraseology such as *data*

[19] Forbes, Why Big Data is the new natural resource, http://www.forbes.com/sites/ibm/2014/06/30/why-big-data-is-the-new-natural-resource/ (June 30, 2014)

exhaust [20] might be symptomatic of a misplaced focus on acquiring big data to the exclusion of addressing the fundamental and possibly more difficult questions of how analytics enables a firm's vision and strategies.

*Managerial Implications A clear message of this study to the discerning practitioner is the importance of shifting from a techno-centric frame to a purposive frame of reference. Second, it emphasizes the importance of starting with a clear vision of how advanced analytics can add value to a firm's strategy enablement and execution. Third, it suggests a shift in focus from big data per se to a focus on developing the infrastructure and talent to exploit the potential of big data and advanced analytics. Fourth, a unique feature of our proposed interpretive model of IT is that apart from serving as a theorizing framework, it can serve as a tool for the practitioner to assess a firm's bottlenecks in leveraging advanced analytics. Figure 3.1 points to a possible use of the framework to assess the current state of technology infrastructure and analytics talent - two areas which our study identified as potential bottlenecks.Thus using the framework in Figure 3.1, a firm or an industry can conduct a similar assessment. Such an assessment we believe will be a useful sense-making exercise which might reveal to the firm or industry the areas that are solved problems in its specific case and the areas where it needs to focus its attention and energies. In sum, we believe that the interpretive model apart from being a useful theorizing framework, can also be fruitfully employed by the practitioner.

[20]Information Week, Big data has exhaust problem,http://www.informationweek.com/big-data/big-data-analytics/big-data-has-exhaust-problem/d/d-id/1278765 (Jun 24, 2014)

Finally, we wish to make a case for IT industry analyst firms such as Gartner to promote a *systems of systems* view of IT similar to the interpretive model of IT. We believe that a coverage of individual technology domains in isolation such as Enterprise Resource Planning(ERP), Manufacturing systems, Storage Systems, Consumer Goods Systems, Supply chain systems , Big Data, Data Architecture etc. has at least the following drawbacks from a practitioner's point of view.
First, it has the effect of all stakeholders such as software vendors, implementation partners and managers within firms being unintentionally complicit in hyping up the flavor of the time, such as an ERP systems, as a panacea. This limited view is probably partly responsible for the infamous trough of disillusionment in Gartner's hype cycles of technologies. A wider system of systems view allows all to see the ERP, despite its scale, as but one component of a wider meta-system of technological affordances. We contend that such a view can help the IT manager reframe the trough of disillusionment as but a natural and dynamic flow of bottlenecks between components technologies as they play catch up with each other in a continual game of adaptive stretch[Arthur, 2009].
Applying this perspective to the case of analytics for instance, we can see that initially, scanning of data was limited to structured RDBMS style data sets entered by humans and hence was the bottleneck. With the advent of sensor networks and IoT, data-glut has shifted the bottleneck to storage of data. Cloud storage technologies seem to have solved that problem but in turn gave birth to another bottleneck - limited bandwidth relative to the volumes of data in the age of big data. A solution that is being offered is to reduce data movement by moving compute logic to the

edges of the network where the data is generated and stored (edge computing). Second, we believe that this tools and technology centric view is partly responsible for a perpetuating an IT community whose denizens receive training in a specific technology which becomes their default lens even as they move up the hierarchy into senior leadership roles. This narrow lens at the top is detrimental to firm as it is akin to firms seeing the disparate parts of the elephant which might in turn result in firms making IT investments driven by hype cycles rather than business needs. We believe, that promoting the systems of systems view such as the interpretive model of IT will help increase the proportion of senior IT leadership that can maintain a meta view of technology, make IT investment decisions based on a clear understanding of its value and hopefully, within their spheres of influence, replace a cyclical disillusionment with technology with hope. Thus, a seasoned IT leader with the system of systems perspective will likely not sell edge computing to her organization as a panacea and can ideally also foresee the locus of the next bottleneck.

3.8 Conclusion and Future work

As the cognitive computing era and its business world progeny, advanced analytics, come into their own, they are likely to push the frontiers of value creation potential of organizations and individuals. In attempting to provide an organizing framework and a vocabulary for discussing this paradigmatic shift, we believe, we have broken new ground by providing a first articulation and delineation of the cognitive computing era in IS literature. We believe that this new lens can be gainfully employed to parse extant IS literature, in the process, identifying opportunities for

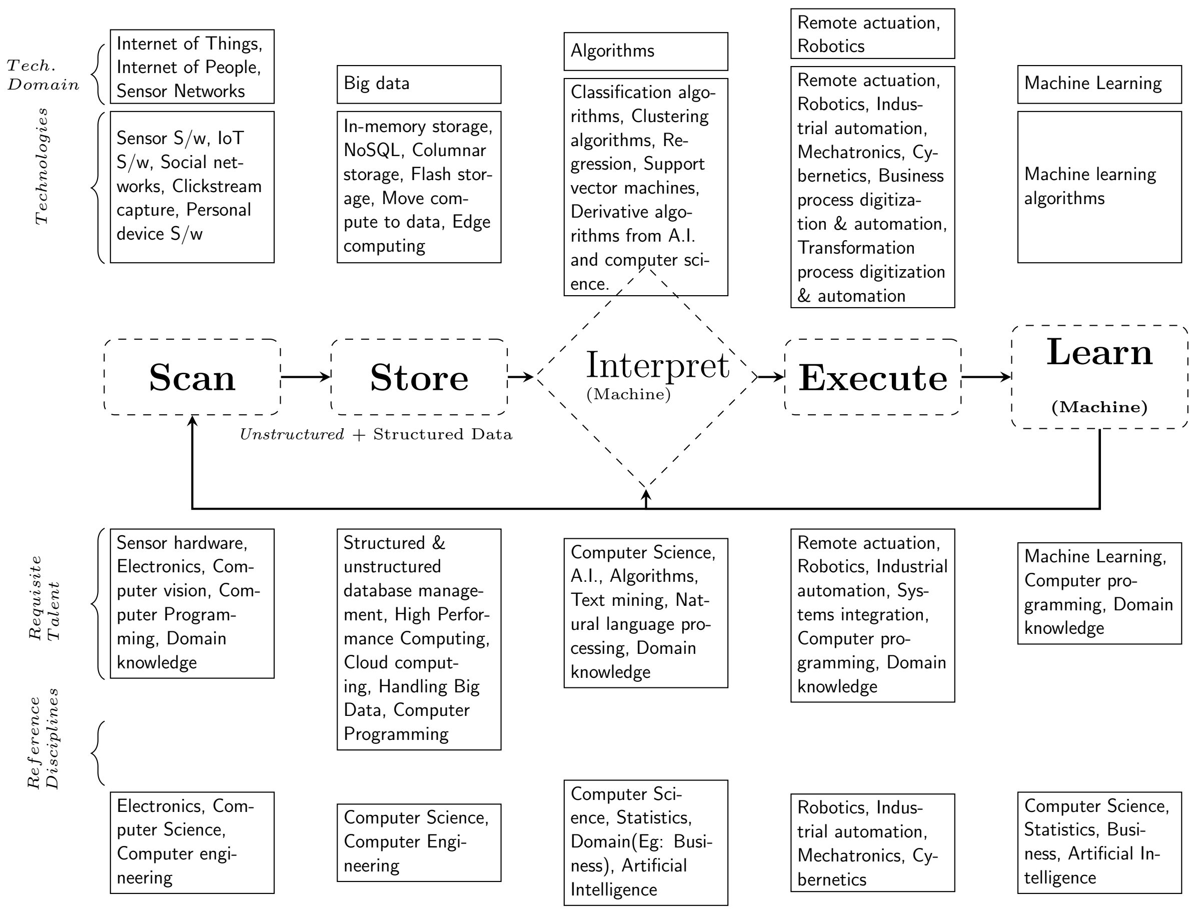


Figure 3.1: Elements of Advanced Analytics parsed through the *Interpretive Model of IT* lens

fruitful research on the cognitive computing era, its applications in business, its business value and its limitations.

Secondly, in deconstructing IT and identifying its constituent parts, we have provided an organizing framework in which extant technologies can be sorted. We believe that if the IS field is to ground itself firmly in the shifting sands of technology, a change in dimensions of the frame of reference from tools and technologies to affordances is essential. Our proposed framework achieves this shift by providing purposive affordances of technology - scan, store, present, interpret, execute and learn. In our future work, we intend to test the robustness, exhaustiveness and stability of this framing by organizing various emerging technologies within this framework. Figure 3.1 gives a flavor of the contours of a possible attempt.

A third direction we intend to take our theorizing is in docking the interpretive model of IT with adjacent reference disciplines to identify opportunities for generating new insights. For instance, we want to indulge our hunch that cognitive computing might fundamentally change the way strategy is done in organizations. Strategy is to a large part making moves for the future much like in chess. If the moves of the opponent or nature are known in advance, it is bound to change the rules of the game. Thus we intend to explore possible new frameworks at the intersection of extant strategy frameworks and the interpretive model of IT.

In sum we feel that in beginning to theorize the cognitive computing era, we have only scratched the surface of this emerging paradigm shift. Much theoretical work needs to be done to deepen our understanding. Possible avenues for achieving a structural deepening of this framework are bibliographic studies of IS literature through the lens of this framework as well as situating the interpretive model of IT within extant IS theoretical frameworks. This is the intent of our future theorizing effort. In sum, we hope that the interpretive model of IT as well as our theoretical framework of man-machine interpretation modes will serve to guide future research on the role of cognitive computing and advanced analytics in the business domain.

On the empirical front, we believe this paper makes some important contributions. To our knowledge, this paper is a first empirical study in IS literature to study traditional and advanced analytics as separate constructs. Further, this study to our knowledge is a first to investigate the differential impact of covariates on the propensity to adopt traditional versus advanced analytics. Third, this paper helps shift the focus from big data per se to a focus on the ability to exploit big data. Thus, it clarifies that in this big data age, data might eventually become as pervasive as air and can barely be considered a valuable, rare and in-imitable resource. It is the organizational lung capacity to separate out the oxygen of insights from the air of big data, that will be key. Some constituents of this lung capacity that our study has revealed are a firm's clarity about business value of analytics, having the right strategic orientation, and possessing the right analytics talent. Our findings also point to a potential path dependence between traditional and advanced analytics capabilities within firms. Empirically, testing the path dependence has not been possible in this paper and is the subject of planned future work.
Additionally, we see frontier territory for empirical studies on the business value of advanced analytics. For instance, we are currently exploring the implications of advanced analytics in reducing inventory write offs - a significant source of value

destruction in the economy. We also believe, that advanced analytics is a technology with significant implications at the individual level. We are currently exploring the role of advanced analytics in amplifying human potential through applications such as learning analytics and career analytics. There is also open territory in exploring the use of advanced analytics in enabling better Government and Governmental services.

A limitation of our study is that it does not address the question firms considering analytics for the first time might have: Should we first develop traditional analytics and then consider advanced analytics? Further, this study is cross-sectional and hence the findings are associational rather than causal in nature. We are also limited in our range of variables due to the secondary nature of our dataset. Our analysis may not have included all potential variables that might affect advanced analytics usage. For instance, future research might explore how senior management attitudes towards business analytics affect its adoption in firms. Thirdly, though InformationWeek randomly selects the respondents, the data is not from a pure random sample and this might limit the generalizability of our finding.

In conclusion, we believe that the paradigmatic shift from programmable to cognitive computing will in turn lead to a quantum and qualitative shift in the way technology enables value creation in individuals, firms, society and Government. It has been our endeavor in this paper to make a first attempt to understand and explicate these shifts and we aspire to continue this journey in our future work.

3.9 Contributions and Limitations

Empirics Contribution To our knowledge, our empirical study is a first to study traditional and advanced analytics as separate constructs. Specifically, this paper empirically tests the antecedents of traditional and advanced analytics. A contribution of the paper is an allusion to a potential path dependence between traditional and advanced analytics capabilities within firms. Empirically, testing the path dependence has not been possible in this paper and is the subject of planned future work. The paper also disabuses us of the perception that once big data is acquired the insights will start flowing in. On the contrary, our study reveals that big data might eventually become as pervasive as air and can barely be considered a valuable, rare and in-imitable resource. Organizations need to develop the lung capacity to separate out the oxygen of insights from the air of big data. Some of the determinants of this lung capacity that our study has revealed are firm's clarity of business value of analytics, having the right strategic orientation, and possessing the right analytics talent.

A limitation of our study is that it does not address the question firms considering analytics for the first time might have: Should we first develop traditional analytics and then consider advanced analytics? Further, this study is cross-sectional and hence the findings are associational rather than causal in nature. Secondly, we are limited in our range of variables due to the secondary nature of our dataset. Our analysis may not have included all potential variables that might affect IT usage. For instance, future research might explore how senior management attitudes towards business analytics affect its adoption in firms. Thirdly, though InformationWeek randomly selects the respondents, the data is not from a pure random sample and this might lint the generalizability of our finding.

Theory Contribution This paper contributes by providing a theoretical framework for IT use in organizations which integrates and delineates both the programmable and cognitive computing paradigms. To our knowledge this is a first attempt at achieving this. We present this delineation of IT into its interpretive and execution enabling roles as a means to advance theorizing and empirics about the role of cognitive technologies in the business domain.

3.10 Appendix

Confirmatory Factor Analysis

Table 3.7: Two-Factor CFA model for Nature of Business Intelligence and Analytics Usage

Measures	Factors	
	BIpred	BItrdn
Business activity monitoring	-	.483*** (.067)
Competitive Analysis	-	.472*** (.067)
Corporate Governance	-	.472*** (.068)
CRM	-	.623*** (.058)
Financial Analysis	-	.455*** (.068)
Fraud Prevention	.585*** (.063)	-
Predictive Analysis	.502*** (.069)	-
Product Development	-	.423*** (.071)
Risk Management	.719*** (.059)	-
Sales Tracking	-	.646*** (.056)
Social Media Sentiment Analysis	.42*** (.072)	-

RMSEA = 0.042 ; SRMR = 0.05; CFI = 0.956
*** $p \leq 0.001$ ** $p \leq 0.01$ * $p \leq 0.05$. $p \leq 0.1$

Table 3.8: Three-Factor CFA model for Managerial Challenges in Implementing Business Intelligence and Analytics

Measures	Factors		
	TalentGap	CompIssue	ValClarityGap
Challenges Getting Agreement on Standardized Products	-	.334** (.111)	-
Ease-of-use Challenges with Complex Software	.239** (.086)	-	-
Integration/Compatibility Issues With Existing Platforms	-	.303** (.834)	-
Lack of Industry Standards	-	.378*** (.11)	-
Lower-than-expected Analytic Value	-	-	.65*** (.129)
No Clear ROI	-	-	.329*** (.108)
No Need for BI Capabilities Throughout our Enterprise	-	-	.288** (.094)
Overlap With Other Products	-	.418*** (.114)	-
Talent Is Too Scarce or Expensive to Hire	.537*** (.108)	-	-
Training Internal Staff Too Time-Intensive and Costly	.766*** (.138)	-	-

RMSEA = 0.01 ; SRMR = 0.044; CFI = 0.993
*** p <0.001, ** p <0.01

Table 3.9: Single-Factor CFA model for Past Orientation of Business Needs

Measures	OrientPast
Optimize Business Operations	.497* (.229)
Marketing Segmentation	.443* (.207)
Maintain Compliance with Laws and Regulations	.217. (.118)

RMSEA = 0.0 ; SRMR = 0.0; CFI = 1.0
* p <0.05, . p<0.1

Table 3.10: Single-Factor CFA model for Future Orientation of Business Needs

Measures	OrientFutr
Predict Promising New Business Opportunities	.592* (.083)
Identify Business Risk	.303*** (.086)
Predict Customer Behavior	.471*** (.079)
Predict Fraud or Financial Risk	.211. (.089)
Predict Sales	.661*** (.084)

RMSEA = 0.078 ; SRMR = 0.047; CFI = 0.92
*** p <0.0001, . p<0.1

Table 3.11: Single-Factor CFA model for Big Data Sources

Measures	BDSources
Analyzing high-scale machine data from sensors, web logs etc.	.375*** (.091)
Analyzing social network comments for consumer sentiment	.609*** (.112)
Analyzing web clickstreams	.571*** (.109)
Finding correlations across multiple, disparate data sources (clickstreams, geospatial, transactions, etc.)	.198* (.094)

RMSEA = 0.0 ; SRMR = 0.007; CFI = 1.0
*** p <0.0001, . p<0.1

Table 3.12: Two-Factor CFA model for Data Analytics Infrastructure

Measures	Factors	
	BDinfra	TDinfra
Complex Event Processing Technology	.652*** (.079)	-
Data Cleansing and Quality Tools	-	.563*** (.063)
Data Integration Software (ETL)	-	.796*** (.052)
Document Imaging and Capture Technologies	-	.365*** (.073)
Hadoop	.534 (.078)	-
High-scale Data Mart/Warehouse	-	.479*** (.068)
NoSQL Database(s)	.46*** (.083)	-
On-Premise Data Mart/Warehouse(s)	-	.607*** (.059)
Trickle Feed/Change Data Capture Technologies	.414*** (.083)	-

RMSEA = 0.067 ; SRMR = 0.056; CFI = 0.911
*** p <0.0001

Table 3.13: Two-Factor CFA model for Data Management Challenges

Measures	Factors	
	BDMgtChlng	TDMgtChlng
Accessing Relevant, Timely or Reliable Data	.452*** (.087)	-
Accessing Content Like Word Files, Emails and Presentations	.348 (.091)	-
Cleansing, De-Duping or Ensuring Consistent Data	-	.391*** (.089)
Coping with Rapidly Increasing Volumes of Data or Content	.361 (.089)	-
Extracting Data/Transactional Info From Paper Forms /Documents	-	.216*** (.093)
Integrating Data (e.g. extract, transform, load or data federation)	-	.469*** (.087)
Maintaining Reliable and Responsive Data Marts/Warehouses	-	.415*** (.087)
Organizing and Maintaining Data Models and Taxonomies	-	.546*** (.086)
Processing High Velocity Data Streams	.548*** (.092)	-
Reducing Data Latency and Supporting Faster Decision Making	.387*** (.089)	-

RMSEA = 0.031 ; SRMR = 0.05; CFI = 0.946
*** p <0.0001

CHAPTER IV

Novel or Feasible? Betting on Technological Innovations in An Open Innovation Funnel: An Empirical Analysis of Idea Selection

Technology is enabling a variety of open innovation models, allowing firms to harness the wisdom of the crowds. A popular typology is firms conducting open innovation contests to collect ideas from which they pursue a few. Since ideas in open innovation contests are sourced from outside the four walls of a firm, they are likely to be endowed with a fresh perspective and be innovative. However, it is not clear whether the open innovation funnel has a propensity to pick innovative or incremental ideas from the superset of ideas that enter it. We explore this question by studying the idea selection process in a multi-stage-gate open innovation funnel of a large private bank, widely recognized as a technology pioneer. Interestingly, we find evidence for a 'stage effect' in the affinity for innovativeness. In the earlier stages of the innovation funnel, decision makers tend to pick innovative ideas but in the later stages seem to prefer incremental ideas. We find an explanation for this surprising volte-face in Construal Level Theories (CLT). Decision makers closer to the point of commitment(of resources) are more conservative than those psychologically and temporally removed. We also find explanations for some of our findings in theories of choice and decision making in the behavioral economics tradition. Managerial implications for firms aspiring to be innovative are presented.

4.1 Introduction

A recent survey of 125 executives of large firms in Europe and United States with annual sales in excess of USD250 million found that 78% of the firms are practicing open innovation [Chesbrough and Brunswicker, 2014]. More tellingly, 82% of the practicing firms reported an increase in intensity of open innovation. This increasing adoption of open innovation by firms is being driven primarily by a desire to have access to more ideas. An implicit assumption is that more ideas will lead to more innovation. However, history is replete with examples of firms failing to leverage technologies developed in their backyards which ended up disrupting their industries [Garud et al., 1997]. Xerox [Smith and Alexander, 1988], Kodak [Lucas Jr and Goh, 2009] and IBM [Langlois, 1997] are but cliched poster children for the many established firms that have failed in the past to pursue innovations within the realm of their awareness and absorptive capacity [Glasmeier, 1991, Henderson, 1993]. In these cases, missing the boat of innovation was not an ideas problem. It was a failure to bet on the right ideas.

Though this puzzling behavior has been chronicled by several scholars [Christensen, 1997, Ghemawat, 1991, Henderson, 1993, Utterback, 1994, Henderson and Clark, 1990], surprisingly, there is little research on what causes established and *innovative* firms to systematically ignore good ideas over long periods of time. More specific to our study, it is not clear whether the relatively recent model of open innovation ameliorates this problem of idea selection. In this paper we attempt to answer this question by empirically studying the idea selection process within a multi-stage-gate open-innovation funnel of a large private bank considered a pioneer in innovative use of technology for banking. We find evidence for early stages of the funnel favoring innovative ideas with the later stages regressing to incremental ideas. We find explanations for this surprising reversal in preference for innovativeness in construal level theories(CLT). In the early stage gates of the open innovation funnel there is a desire to be innovative and unique. However, as we progress through the funnel, the contingencies of execution and feasibility seem to regress the funnel to the status quo of sustaining innovations. We explain our results using theories in behavioral economics such as prospect theory, status quo bias and availability bias. Managerial implications of our findings are discussed.

To our knowledge, by focusing on idea selection rather than idea generation, we are breaking new ground in the IS open innovation literature. In an effort to highlight our focus on idea selection, we start out in Section 4.2 by articulating a decision theoretic model of the innovation funnel. This model at once delineates and integrates the idea generation and idea selection functions of the innovation funnel. Next, we review IS literature on innovation to situate our work within extant work on innovation. In Section 4.3 we draw on theories from behavioral economics to posit our hypotheses about idea selection in an innovation funnel. Our research setting is described in Section 4.4. Measures are described in Section 4.5 and Data is presented in Section 4.6. Our empirical model is covered in Section 4.7 and results discussed in Section 4.8. We draw our conclusions in Section 4.9. We close out with the contributions and limitations of our study in Section 4.10.

4.2 Literature Review

A decision theoretic framework of innovation funnels Relative to a long history of academic discourse on detailed aspects of the new product development process [Krishnan and Ulrich, 2001], innovation funnels, responsible for feeding new ideas into the product development process have only recently gained attention as an area of inquiry. Surprisingly, despite their centrality in the innovation process, they have been typically treated as black boxes. We depart from this approach by opening up the innovation funnel apparatus and recasting its innards in the mold of Daft and Weick's scanning, interpretation and learning framework [Daft and Weick, 1984]. In building a model of organizations as interpretation systems, Daft and Weick (1984) liken firms to a central nervous system of sorts which obtains, filters and processes information to make choices. Seen through this theoretical lens, innovation funnels [Chesbrough, 2003], can be framed as an organizational tool for idea collection and idea selection, with idea implementation being a part of new product development (NPD). Thus, the scanning phase can be considered as seeking out ideas. Interpretation involves analyzing the idea on explicit or implicit dimensions that the organization considers important and making the choice to either pursue or drop the idea. Learning occurs in the act of implementing the chosen ideas. This functional decomposition of the innovation funnel sets the context for our study (Figure 4.1) with our focus being on the idea selection phase.

Open innovation literature review Langlois[Langlois, 2003] has argued that managers must find new ways to conceptualize the post-Chandlerian firm where innovation proceeds along less hierarchical lines because large vertical integrated organi-

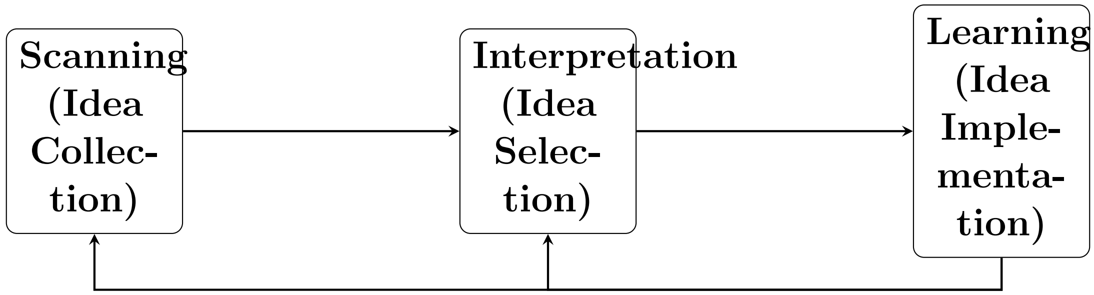


Figure 4.1: Innovation Funnels as an Idea Interpretation Tool
(Source:Daft and Wieck, 1984)

zations are becoming less significant and are joining a richer mix of organizational forms [Langlois, 2003], p. 353. This call for a more federated innovation process is consistent with a rise in the need to access external sources of knowledge [Chiesa and Manzini, 1998, Haour, 1992, Narula and Hagedoorn, 1999]. Most end products embody an increasingly diverse range of technologies (e.g., mobile phones), each of which requires specialized domain expertise to develop. Thus, firms that adhere to the Chandlerian model of vertically integrated innovation are slowly becoming archaic [Iansiti, 1997]. By one account, on average, external sources account for between 35% and 65% of the inputs important to the development of successful innovations [Conway, 1995]. In the past as well, there has been a long tradition of firms partnering for knowledge creation. Such partnerships have included joint ventures and alliances (Hamel, 1991; Inkpen, 1992; Kogut, 1988; Mowery, Oxley, & Silverman, 1996), suppliers and customers (Hakanson & Johanson, 1992; Von Hippel, 1988), federation of competitors [Chiesa and Manzini, 1998, Hagedoorn, 1993, Ingham and Mothe, 1998] and university collaborations [Bailetti and Callahan, 1992, Conway, 1995]. A common theme in these traditional forms of coming together, is a strong coupling between the partners. A clear definition of the relationship, usually achieved by legally binding clauses, precedes work on knowledge creation.

Open innovation departs from this precedence on relationship structuring to an almost singular focus on innovating the artifact. We contend that Linux, and the open source software movement it helped birth, marks this point of departure. The model of open source software development has since proliferated into a wide variety of open innovation phenotypes such as innovation platforms, contests and the like. Till recently, the artifact being innovated on these platforms have been information goods. However, with the increasing convergence between bits and atoms, manifest in developments such as 3D printing, the physical world is also opening up to being 'open-innovated'. For instance in June, 2013, GE launched a 3D printing design quest to redesign a metal jet engine bracket making it 30 percent lighter while preserving its integrity and mechanical properties like stiffness. The winning design from among 700 submissions from 56 countries, came from an engineer in Indonesia and achieved an 84 percent reduction in weight.

Cases such as these raise the exciting prospect of open innovation affecting a shift towards a *democracy* in idea generation and a *meritocracy* in idea selection. The idea generation aspect of open innovation funnels has been increasingly studied in recent years ranging from prescriptions for designing innovation funnels [Cooper and Kleinschmidt, 1993, Ding and Eliashberg, 2002], determinants of idea success [Johne and Snelson, 1988] and means to improve the quality [Girotra et al., 2010, Kornish and Ulrich, 2011, Reitzig, 2011] of ideas entering the innovation funnel. The idea selection aspect of the innovation funnel on the other hand is understudied [Blair and Mumford, 2007], probably, in part due to an assumption that generation of more and better quality ideas will lead to better innovation [Osborn, 1953, Diehl and Stroebe, 1987, Gallupe et al., 1992]. This might be due to a naive belief in an inherent meritocracy in idea selection, wherein the best ideas get selected. However, not all innovations concern well-defined engineering problems with clear criterion for selection. In fact the *fuzzy front end* of new product development is well documented[Smith and Reinertsen, 1991]. Additionally, as we move from screening explicit and structured incremental innovations [Booz, 1982, Urban,], such as bracket design, to discontinuous innovations, the fuzziness of decisions increases. Further, with an increase in the number of ideas[Payne, 1982, Billings and Marcus, 1983] in the funnel and the number of attributes under consideration[Malhotra, 1982, Shields, 1983], the complexity of evaluating ideas quickly increases. An increase in the complexity of decision making is associated with an increase in cognitive costs. In an effort to reduce cognitive costs, decision makers might resort to mental shortcuts in selecting ideas. Thus, despite the best intentions to select the best ideas, innovation funnels might be bounded in their rationality. Indeed, we contend that these mental shortcuts and heuristics in decision making might hold candidate explanations for the paradox of innovative firms such as Kodak turning a blind eye to industry redefining innovations. This belief is our motivation for drawing on theories from behavioral economics and theories of choice and decision making under uncertainty, to develop our hypotheses about the idea selection process in an open innovation funnel.

4.3 Theory and Hypotheses

Decision strategies in an innovation funnel: Exclusion Vs. Inclusion An innovation funnel consists of a set of stage gates through which ideas pass to eventually enter the pipeline of new product development [Chesbrough et al., 2006]. Each stage gate can be considered as an organizational interpretation lens [Daft and Weick, 1984] that evaluates an idea on a set of attributes and lets it through or screens it out. In this sense, information processing at each stage gate can be typified as deciding to choose from a multi-attribute consideration set under uncertainty (of innovation success).

Decision-making strategies can be classified under two broad categories: compensatory and non-compensatory. Compensatory decision strategies make tradeoffs between choice attributes wherein a good value on one attribute can make up for bad values on other attributes for a choice. In Non-compensatory strategies on the other hand, a bad value on a critical attribute will ensure that a choice is not selected no matter how good it is on other attributes. The essence of the difference between these two strategies is how they deal with conflict. Compensatory rules confront conflict, whereas non-compensatory rules avoid it. Non-compensatory rules require much less cognitive effort due to smaller number of mental operations and are also known as exclusion or *elimination* strategies. Compensatory decision rules on the other hand require a larger number of mental operations, are cognitively costlier and also known as inclusion or *selection* strategies.

Decision making theory has established that cognitive effort(attention) is the scarce resource in decision making [Simon, 1978]. All else being equal, decision makers prefer more accurate choices and less effortful choices. As cognitive costs increase, decision makers tend to choose elimination decision strategies like elimination by aspects (EBA)[Tversky, 1972], satisfcing [Simon, 1972] and lexicographic heuristics in an attempt to reduce cognitive load. Additionally, there is evidence that people find making explicit tradeoffs emotionally uncomfortable [Hogarth, 1987] . Thus, decision makers may avoid exhaustive assessment of all attributes, not only because they are difficult to execute (cognitive effort) but also because they require the explicit resolution of difficult value tradeoffs (conflicts)[Gregory et al., 1991]. In an experimental study investigating strategies for narrowing choice options, 70% of the participants preferred elimination over selection strategies [Heller et al., 2002]. Indeed, a review of contemporary stage gate systems currently in use in established firms finds evidence for existence of *must-meet* criteria to weed out misfit projects quickly followed by a more exhaustive scoring on *should-meet* criteria to prioritize projects [Cooper, 2008]. In general, research has provided evidence for the primacy of elimination strategies as a means to reduce the consideration set, especially as the number of attributes and alternatives increases [Ford et al., 1989, Payne et al., 1993]. More interestingly, it has been found that elimination strategies despite having considerably less cognitive costs, in some situations, provide choices of equivalent accuracy[Payne et al., 1993].

Thus, we contend that in an innovation funnel elimination strategies will tend to precede selection strategies (Figure 4.2), especially as the number of alternatives and attributes increases. The earlier stage gates are likely to use elimination choice strategies to reduce the consideration set while the later stage gates are likely to engage in a more exhaustive evaluation of ideas. To summarize, we put forth the following propositions which is also illustrated through Figure 4.2

Proposition 1. Cognitively less costly decision strategies are more likely to be employed in earlier stage-gates of innovation funnels.

Proposition 2. Cognitively more costly decision strategies are more likely to be employed in later stage-gates of innovation funnels.

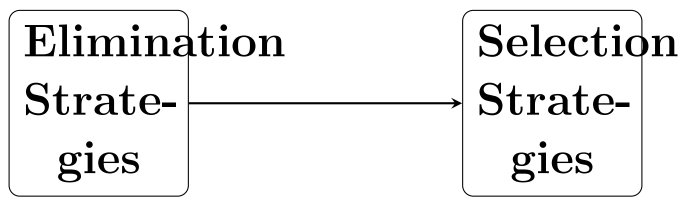


Figure 4.2: The precedence of Elimination over Selection choice strategies

Attributes for selecting ideas in an innovation funnel Innovations are evaluated for alignment within contexts in which they are situated [Leonard-Barton, 1988]. Kuan and Chau [Kuan and Chau, 2001] have identified three contexts salient to adoption of technological artifacts - technology , organization and environment (TOE, Figure 4.3) . Extant literature has identified aspects of each of these contextual layers with which the attributes of an innovation need to align for a successful outcome [Leonard-Barton, 1988].

At the environment level, the innovation needs to align with the socio-cultural norms of the market, be compliant with regulations, and, be compatible with industry

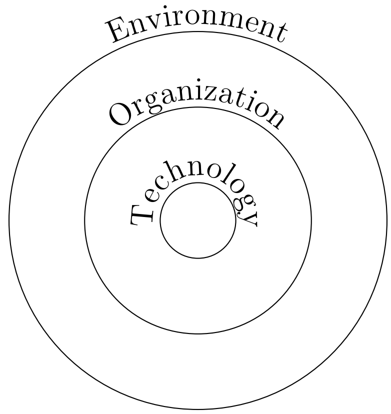


Figure 4.3: The Contexts of an Idea

standards and infrastructure[Adner and Kapoor, 2010]. At the organization level, the technological innovation needs to align with the organization's culture, identity, vision and resources. At the technological innovation level, extant literature is replete with a wide variety of attributes on which innovations may be evaluated. In an exhaustive examination of 90 studies on creativity and idea generation, Dean et. al. [Dean et al., 2006] have identified four dimensions that are regularly used for evaluating ideas - novelty, workability, relevance and specificity . Within each of these dimensions, the authors identify two measurable sub-dimensions for each of these dimensions (Table 4.1).

In what follows, we analyze different factors, in each of the TOE contextual layers, starting with the outer layers first. We contemplate each factor's salience for idea selection in innovation funnels, in light of theories in behavioral economics, and theories of decision making and choice. We complement our theoretical understanding with findings from previous empirical studies. Based on our analyses we propose our hypotheses.

Table 4.1: Definitions of Idea Quality Dimensions and Sub-Dimensions

fs	Range	Definition
1	**Novelty**	The degree to which an idea is original and modifies a paradigm
1.1	Originality	The degree to which the idea is not only rare but is also ingenious, imaginative, or surprising
1.2	Paradigm relatedness	The degree to which an idea is paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational
2	**Workability (Feasibility)**	An idea is workable (feasible) if it can be easily implemented and does not violate known constraints
2.1	Acceptability	The degree to which the idea is socially, legally, or politically acceptable
2.2	Implementability	The degree to which the idea can be easily implemented
3	**Relevance**	The idea applies to the stated problem and will be effective at solving the problem
3.1	Applicability	The degree to which the idea clearly applies to the stated problem
3.2	Effectiveness	The degree to which the idea will solve the problem
4	**Specificity**	An idea is specific if it is clear (worked out in detail)
4.1	Implicational explicitness	The degree to which there is a clear relationship between the recommended action and the expected outcome
4.2	Completeness	The number of independent subcomponents into which the idea can be decomposed, and the breadth of coverage with regard to who, what, where, when, why, and how
4.3	Clarity	The degree to which the idea is clearly communicated with regard to grammar and word usage

Source: [Dean et al., 2006]

4.3.0.1 Environment Attributes: Regulatory constrains and Innovation Ecosystem constraints

Institutional theory contends that any change within an organization is situated in the context of markets and institutions [Greenwood and Hinings, 1996]. However, not all institutional pressures are created equal and the response of the firm might range from passive conformity to proactive manipulation[Oliver, 1991], depending on how permeable each of these institutional layers is perceived [Greenwood and Hinings, 1996]. The inner contextual layers will tend to be assessed within a resource dependence theory mindset where the organization might consider a wide range of active choice behaviors to manipulate and influence the context [Scott, 1987]. By comparison, the peripheral contextual layers will likely be seen through an institutional theoretic lens which tends to limit firm options to different types of structural or procedural conformity to the environment (ibid).

Two such factors that trigger preconscious acceptance and conformity rather than active resistance or manipulation are regulatory constraints [DiMaggio and Powell, 1983, Zucker, 1977] and innovation ecosystem constraints[Adner and Kapoor, 2010]. Innovation ecosystem is one of the most comprehensive articulation of the environment in which innovations must operate[Adner, 2006, **?**] . The key idea is that innovations rather than succeeding in isolation must thrive with complementary offerings from other players constituting the ecosystem[Adner, 2012]. Innovations that do not align with the existing innovation ecosystem, face a high probability of failure. Indeed a a key strategy suggested for innovations facing integration and interdependence risks in the innovation ecosystem is to consider delaying product development to let other players catch up[Adner, 2006]. In sum, both regulatory constraints and innovation ecosystem constraints are factors that the firm is likely to perceive out of its span of control. Consequently, they are likely to serve as an easy *fate-accompli* heuristic for the decision makers to eliminate ideas. Thus we hypothesize:

Hypothesis 6. Innovations that face regulatory constraints are likely to be eliminated from innovation funnels.

Hypothesis 7. Innovations that face innovation-ecosystem constraints are likely to be eliminated from innovation funnels.

4.3.0.2 Technology Attribute: Innovation Uniqueness

There is consensus that novel solutions are generally characterized as being new and useful[Amabile, 1996, Mayer, 1999, Niu and Sternberg, 2001, Plucker et al., 2004]. Though there is no universal definition of uniqueness, one plausible candidate is - an idea is unique if it has not been expressed before[MacCrimmon and Wagner, 1994]. An analogous trait is originality, which refers to the idea being surprising, imaginative, uncommon or unexpected [Ang and Low, 2000, Dean et al., 2006]. The evidence from product development literature on whether uniqueness is correlated with a product being selected or rejected for development is mixed [Cooper and Kleinschmidt, 1990, More, 1982]. Some studies have found support for unique ideas being rejected which has been explained by the greater uncertainty and thus risk such ideas present to the firm [More, 1982]. Prospect theory [Tversky and Kahneman, 1991, Tversky and Kahneman, 1991] contends that one of the basic phenomenon of choice under both risk and uncertainty is that losses loom larger than gains. Hence, when presented with choices that are unfamiliar, uncertain and appear risky, decision makers will tend to avoid them. On the other hand, some scholars have found support for innovative products being less likely to be killed [Schmidt and Calantone, 1998]. The justification has been that innovative products often present great opportunities for firms wishing to extend their activities into new markets and technologies, involving greater emotional and strategic commitment, and may therefore be tougher to terminate [Danneels, 1998, Schmidt and Calantone, 1998]. Since one of the primary purposes of open innovation initiatives is to source novel ideas from outside organizational boundaries [Chesbrough, 2003], we hypothesize that:

Hypothesis 8. Innovations rated higher on uniqueness are more likely to be selected in open innovation funnels.

4.3.0.3 Technology Attribute: Incremental Innovation

Another attribute of novelty is paradigm relatedness[Besemer and O'QUIN, 1986, Finke et al., 1992, Nagasundaram and Bostrom, 1994]. This refers to an idea′s transformational character, and describes the degree to which an idea helps to overcome established structures, i.e., how radical or revolutionary it is [Besemer and O'QUIN, 1986, Christiaans, 2002]. Ideas that are in line with the current structures are considered sustaining or incremental. Ideas that go against the grain are considered radical. Radicalness of an innovation can be defined as the extent to which the technology advances the performance frontier faster than the existing technological trajectory

[Gatignon et al., 2002]. Radical products, on the other hand, are more likely to be killed with the justification that radical products are more risky and less clearly tied to market needs and thus receive less internal organizational support[Green et al., 1995] . Further, there is empirical evidence for radical innovations requiring more resources to successfully commercialize[Ettlie and Rubenstein, 1987]. As Schoonhoven et al. [Schoonhoven et al., 1990] point out, technological innovations that require the firm to create new knowledge as part of the innovation process represent a substantial challenge. Foster [Foster, 1986] agrees stating that the major cultural difficulty firms face in managing through technological discontinuities is making skill transitions. Thus more radical innovations have been depicted as creating greater knowledge demands for the company [Dewar and Dutton, 1986, Nord and Tucker, 1987].

On the other hand, research has found that that innovations building on existing competencies are more rapidly introduced and are positively associated with commercial success, particularly when they are incremental[Gatignon et al., 2002].The evaluation of products situated in familiar environments, i.e., targeted at a familiar market or using a familiar technology, benefit from clearer signals regarding potential success. In addition, in introducing the availability effect,Tversky and Kahneman [Tversky and Kahneman, 1973] assert that images of the future are shaped by experiences of the past. In describing this decision heuristic, they established that when faced with the task of judging and making choices, people evaluate the probability of an event by its availability, i.e. by the ease with which particular choices can be recalled or associated with the familiar. Thus we expect that when faced with choosing between ideas that maintain the current trajectory of innovations versus those that take on a different path, decision makers are likely to choose familiar innovations. Thus all the theoretical cannon seems to suggest that incremental innovations will be favored over radical innovations. However, open innovation funnels are specifically intended to depart from this traditional tyranny of the status quo[Samuelson and Zeckhauser, 1988]. Hence we hypothesize that:

Hypothesis 9. Innovations rated higher on being incremental are more likely to be rejected in open innovation funnels.

4.3.0.4 Technology Attribute: Innovation Feasibility

From the innovator's perspective, an idea's feasibility is another vital dimension of idea quality. This dimension captures the ease with which an idea can be transformed into a commercial product [Kristensson et al., 2004, Soll, 2006]. Feasibility is important because although innovative ideas are typically desirable, they will not be implemented if they are not feasible. Hence, the usual definition of a good' idea is an idea that is both highly original and highly feasible [Diehl and Stroebe, 1987]. The evidence accrued by [Abbey and Dickson, 1983] and [Sternberg and Lubart, 1996] indicates that in addition to considering a number of unique attributes of an idea, decision makers also take into account feasibility and likelihood of success. Experimental studies have identified the strong tendency of participants to select feasible and desirable ideas, at the cost of originality as the main reason for their poor performance on selecting innovative ideas[Rietzschel et al., 2010]. Further, firms prefer ideas that hold promise of scaling across regions as well as business units. In fact, evidence suggests that the decision to terminate a new product is often made under uncertainty regarding its likely technical and commercial feasibility [Balachandra, 1984]. Thus we hypothesize that:

Hypothesis 10. Innovations rated higher on feasibility are more likely to be selected in open innovation funnels.

4.3.0.5 Technology Attribute: Business Value

Though in open innovation contexts idea novelty and radicalness are desirable, there is an underlying need to be valuable to the firm and its stakeholders. Baumann and Martignoni [Baumann and Martignoni, 2011] note how researchers in organizational theory and strategic management fields advocate firms to balance exploration of new ideas (seeking novelty) with exploitation of current avenues (seeking utility). Usefulness is the extent to which the idea responds to or solves a problem that is tangible and vital [Amabile, 1996, Dean et al., 2006]. This dimension is also named as an idea′s value or relevance [Dean et al., 2006, Kristensson et al., 2004, MacCrimmon and Wagner, 1994]. In the scope of new product development, this refers frequently to an idea's financial potential [Cady and Valentine, 1999, Franke and Hienerth, 2006, Lilien et al., 2002, Rochford, 1991, Soll, 2006], the strategic importance in terms of enabling competitive advantages [Cady and Valentine, 1999, Lilien et al., 2002, Rochford, 1991], as well as the customer benefit that an idea endows [Piller and Walcher, 2006]. Thus we hypothesize that:

Hypothesis 11. Innovations rated higher on their business value to the firm are more likely to be selected in innovation funnels.

4.3.0.6 Organization Attribute: Innovation Team potential

While products may have different levels of inherent' creativity and acceptability [Barlow, 2000], the skill of the creator in selling to or persuading evaluators that the product is creative is also a critical element in judgments of new products. [Sternberg et al., 2003] argues that individuals judged to be creative are also those who are likely to be capable at selling their products and ideas. Similarly, [Staw, 2009] theorizes that extremely creative people not only produce creative ideas but are also able to recognize and sell their creative ideas. This idea of salesmanship is also similar to Grifitth et. al.'s idea concretization [Griffiths-Hemans and Grover, 2006]. This salience of team potential is also corroborated in the VC decision-making literature which consistently finds entrepreneurial team capabilities as the most critical factor when investing in ideas [MacMillan et al., 1986, MacMillan et al., 1987, Tyebjee and Bruno, 1984, Timmons et al., 1987]. Thus we hypothesize that:

Hypothesis 12. Innovations proposed by teams perceived to have higher potential, are more likely to be selected in open innovation funnels.

In sum, the central assertion of this paper is that open innovation funnels will achieve their intended goal of bringing in fresh ideas from outside the walls of the firm. Thus, they are expected to select unique and scalable innovations while rejecting incremental innovations. In the following section we describe the setting in which test our hypotheses.

4.4 Research Setting

Our research setting is an open innovation initiative by a large private bank in India This bank has been widely recognized as a pioneer in offering innovative banking solutions and services leveraging new technologies.

The objective of the open innovation program is to leverage the innovation potential of the youth to generate technological innovations for the bank as well as foster entrepreneurship in the country. As part of this program, the bank invites students from 22 prestigious engineering schools in India to pitch ideas in the categories of internet banking, paperless banking, mobile banking, biometric devices, near field banking, tablets, mobile ATMs, green banking, handheld devices and self-service banking among others. The program consists of three main stages - Ideate, Prototype and Be an Entrepreneur. Ideas selected for prototyping are actively mentored by senior executives of the bank in addition to receiving an INR 50,000 cash award. Prototypes selected for commercialization are funded and facilitated for venture creation by the bank with ownership of the startup staying with the team. We contend that because the bank is committing significant resources in terms of senior management time and fund, the innovation funnel in this program is representative of real life innovation funnels operative in large established organizations. *Stage Gates in the Open Innovation Funnel Before an idea is selected for prototype phase, it goes through several stage gates in the innovation funnel (Figure 4.4). A brief description of the various stage gates in the funnel is in order:

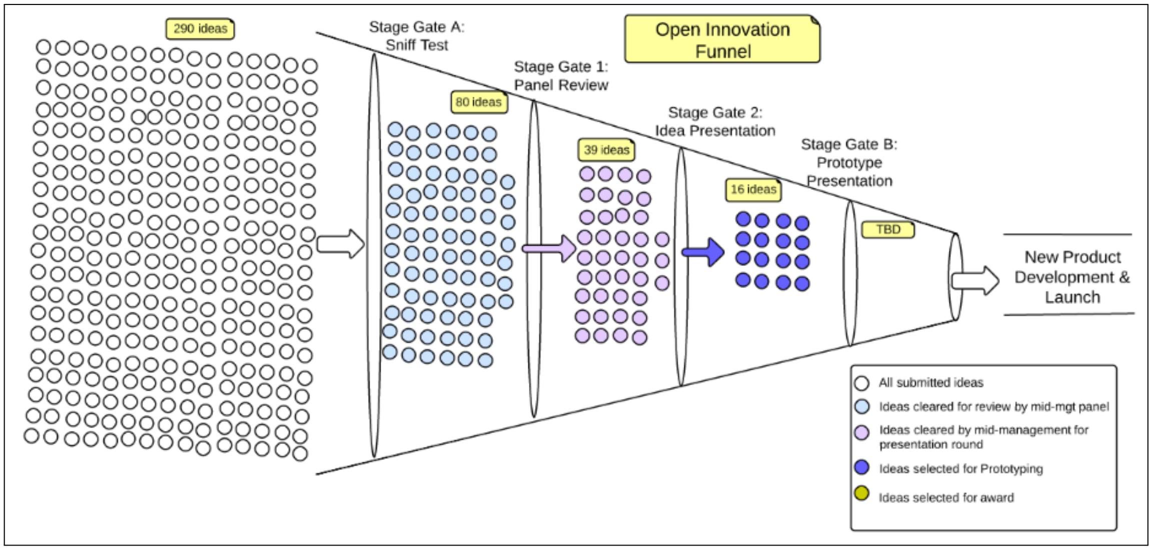


Figure 4.4: Innovation Funnel Stage Gates

Pre Idea-Submission Phase Activities:

The program is orchestrated by the innovation department of the bank that carries out promotion activities on all participating college campuses. Promotion activities consist of banners and posters two weeks prior to launch followed by an on-campus presentation and a question and answer session.

Team registration and concept note submission:

Interested students form teams (max. Of five students per team) and register themselves at the program website. Post registration, the participating teams submit a concept note on their idea.

Stage Gate A (Sniff test)

Initial screening of the ideas by a panel consisting of members from the bank′s dedicated innovation team. At this stage, ideas that are poorly formed, incomplete or are completely misaligned with the bank′s business (ex: enter on line groceries business) are weeded out. Ideas that pass this sniff test, move on to the next stage.

Stage Gate 1 (Panel Review)

The decision makers at this stage are panels of mid-level managers. The panel members read the concept notes submitted by the teams and assign scores to the ideas on five specified attributes (Uniqueness, Scalability, Benefit to bank, Benefit to customer and Technology radicalness). Based on these scores and their judgment, the panel members select the ideas that move on to the next stage.

Stage Gate 2 (Idea Presentation Round)

The decision makers at this stage are a panel of senior-level managers from the bank, external senior level industry veterans and academic experts. The decision makers listen to a 30 minute presentation by the team and follow up with questions to the team to vet the quality and execution potential of the idea and the team. The stakes at this stage gate are higher since ideas that cross this stage gate compete for the scarce senior management mentoring and funding through the prototyping phase. Also, at this stage, the jury looks at both the quality of the ideas as well as the entrepreneurship potential of the team.

Stage Gate B (Prototype Presentation Round)

The decision makers at this stage are senior level managers from the bank. Ideas selected at this stage enter the commercialization phase. At this stage gate the consideration is whether an idea can scale and has good value proposition for the bank.

We do not have data on Stage Gate B yet since the final decisions on which ideas to commercialize have not been taken. Thus Stage Gate B is not part of our study. We also take Stage Gate A out of our analysis since the ideas screened out at this stage are of a relatively poor quality and were removed for a lack of hygiene factors. The focus of our analysis in this paper are thus Stage Gates 1 and 2.

4.5 Measures

Dependent variables In this study we are exploring the impact of TOE factors on the probability of selection of innovations at stage gates 1 and 2. Thus our dependent variables of interest are the binary outcomes at both the stage gates. At Gate 1, the outcome is observed for all ideas that were evaluated. At Gate 2, only the ideas that were selected at Gate 1 are evaluated and hence only their outcome variable is observed. Independent variables The independent variables in our study are informed by the TOE framework. We also draw heavily on the scale developed by Dean et. al. [Dean et al., 2006] for common attributes used for idea evaluation, based on an exhaustive examination of 90 studies on creativity and idea generation. Dean et. al. identified four dimensions that were regularly used for evaluating ideas -

novelty, workability, relevance and specificity. For each of these four dimensions, two measurable sub-dimensions were identified. Confirmatory factor analysis revealed high loadings among the sub-dimensions that comprise each dimension as well as high discriminant validity between dimensions. Further, high inter-rater reliability was achieved even when these dimensions were applied by different raters to different problems. This scale is provided in Table 4.1. We select from this scale the attributes salient to our study.

4.5.0.7 Innovation Uniqueness

The idea quality dimension scale we draw upon (Table 4.1) defines Idea originality as "the degree to which the idea is not only rare but is also ingenious, imaginative, or surprising". The review panel at Gate 1 scored each evaluated innovation on its uniqueness on a scale of 0-30. Innovation uniqueness for purposes of this scoring was defined as "The degree to which proposed innovation is new to the bank, the banking industry or across all industries". Ideas considered new to the bank had a suggested score of 10. Ideas considered new to the banking and financial services industry(BFSI) had a suggested score of 20. Ideas considered completely new and unheard of had a suggested score of 30. The suggested scores were just guidelines and the panel was free to pick any value between 0-30. We had access to this score from the bank for each of the innovations evaluated at Gate 1 and we used it to measure Innovation uniqueness.

4.5.0.8 Innovation Scalability

The idea quality dimension scale we draw upon (Table 4.1) defines Idea implementability as "the degree to which the idea can be easily implemented". The review panel at Gate 1 scored each evaluated innovation on its scalability on a scale of 0-15. Innovation scalability for purposes of this scoring was defined as "The degree to which proposed innovation can be scaled across business units or geographies". Innovations considered applicable only to specific customer segments such as 'wealth management' or 'personal banking' had a suggested score of 8. Innovations considered applicable pan-India for a business vertical such as retail-banking had a suggested score of 10. Innovations considered applicable pan-India across multiple business vertical such as retail-banking and corporate-banking, had a suggested score of 12. Innovations considered applicable pan-India across all the group companies of the bank had a suggested score of 14. Innovations considered applicable across the global operations of the bank had a suggested score of 15. The suggested scores were just guidelines and the panel was free to pick any value between 0-15. We had access to this score from the bank for each of the innovations evaluated at Gate 1 and we used it to measure Innovation scalability.

4.5.0.9 Innovation Business Value

The idea quality dimension scale we draw upon (Table 4.1) defines idea effectiveness as "the degree to which the idea will solve the problem". The review panel at Gate 1 scored each evaluated innovation on its benefit to the bank on a scale of 0-40. Benefit to the bank for purposes of this scoring was defined as "The degree to which innovation will increase firm revenue, efficiency or customer satisfaction". Innovations that are likely to provide a marginal advantage to the bank had a suggested score of 4. Innovations that are likely to provide a single, one time significant advantage to the bank had a suggested score of 20. Innovations that are likely to provide a significant advantage to the bank over the short term had a suggested score of 25. Innovations that are likely to provide a significant advantage to the bank over the medium term had a suggested score of 32. Innovations that are likely to provide a substantial competitive advantage to the bank sustained over the long term had a suggested score of 40. The suggested scores were just guidelines and the panel was free to pick any value between 0-40. We had access to this score from the bank for each of the innovations evaluated at Gate 1 and we used it to measure Innovation business value.

4.5.0.10 Regulatory constraint

Extant literature has identified the conformity inducing influence of regulatory constraints on firm behavior[DiMaggio and Powell, 1983, Zucker, 1977]. The banking and financial services industry is probably one of the most highly regulated industries. Thus, it was not surprising to find mentions of regulatory hurdles in the textual notes and comments provided by the panel for each innovation they evaluated. We had access from the bank to these text comments which included comments on reasons for rejecting ideas, ideas and suggestions on how to improve the innovation, references to other innovations that the bank was working on or had worked on in the past, as well as general notes and comments. With the help of two members from the innovation department at the bank, we closely parsed these text-comments for mentions of regulatory hurdles or constraints. Innovations that had mentions of such concerns, were coded as '1' for Regulatory constraint. Rest were coded as '0'.

4.5.0.11 Ecosystem constraint

One of the key thesis of the recent literature on ecosystem innovation is that innovations rather than succeeding in isolation typically must thrive within an ecosystem of complementary offerings from other players[Adner, 2012]. The traditionally, vertically integrated banking industry is also morphing into an ecosystem consisting of mobile devices, payment services , biometric systems and data services. The innovation contest under consideration mirrors the technology centric nature of this shift. More than a quarter of the submitted innovations were categorized as pure play mobile or internet banking ideas(Table 4.3). Some other categories included biometric systems, internet banking, bluetooth banking and near field communication (NFC) banking. In light of this, a salient aspect of our research setting is a developing country context in which there is likely to be a disparity in the technological evolution of different sub-systems making up the ecosystem[Dedehayir and Mäkinen, 2011]. This was our motivation for parsing the text comments for instances of innovations being held up due to reverse salience in constituent sub-systems of the ecosystem. Some examples are references to the absence of technology in the more than 12 million mom and pop retail shops which make up 95% of the retail industry in India[Srivastava, 2008]. Another example is reference to the lack of a unique identifier for Indian citizens similar to the social security number in the US[Romero, 2012]. We coded innovations that had comments referring to such holdups as '1' for Ecosystem gap. Rest of the innovations were coded as '0' on this measure.

4.5.0.12 Incremental innovation

The idea quality dimension scale we draw upon (Table 4.1) defines paradigm relatedness as "the degree to which an idea is paradigm preserving (PP) or paradigm modifying (PM). PM ideas are sometimes radical or transformational". This refers to an idea?s transformational character, and describes the degree to which an idea helps to overcome established structures, i.e., how radical or revolutionary it is [Besemer and OQUIN, 1986, Christiaans, 2002]. Ideas that are in line with the current structures are considered sustaining or incremental. Ideas that go against the grain are considered radical. The panels reviewing the ideas are senior and mid-level managers who are aware of the different innovation initiatives going in the bank. When they encountered innovations that are similar to those the bank is already working on, they have called them out in their text comments. Most of these comments refer to three broad categories of projects: Projects the bank is currently working on, has worked on in the past and abandoned, has considered in the past but decided not to pursue, has considered in the past but placed on the back burner. For instance, an idea for mobile van banking elicited a comment about a previous attempt by the bank. We closely parsed these text-comments for such references and coded them as '1' for the variable 'incremental innovation'. The rest of the innovations were scored as '0' on this measure.

4.5.0.13 Innovation team potential

A unique aspect of this innovation contest is that the bank expects the winning teams to eventually create a business around the innovations that they proposed. Thus, the bank is interested in assessing the team's potential in addition to that of the innovation. The team assessment is achieved at stage gate 2 when the panel has the opportunity to see a presentation and ask the teams questions. Thus, the text comments by the panel at stage gate 2 contain analysis of the quality of the team and its potential. We closely parse the text-comments for qualifiers identify particular teams and coded them as '1' for team potential. The rest of the innovations were coded as '0' on this measure.

A summary table with definitions of these variables is provided in Table 4.2:

Table 4.2: Definitions of Innovation Evaluation Attributes

Innovation TOE Attributes	Range	Definition
Innovation Uniqueness	0-30	The degree to which proposed innovation is new to the bank, the banking industry or across all industries
Innovation Scalability	0-15	The degree to which proposed innovation can be scaled across business units or geographies
Innovation Business Value	0-40	The degree to which innovation will increase firm revenue or efficiency
Incremental Innovation	0(N), 1(Y)	Whether innovation is a sustaining innovation in line with current technological trajectory of bank
Regulatory Constraints	0(N), 1(Y)	Whether the innovation will face regulatory hurdles
Ecosystem Constraints	0(N), 1(Y)	Whether innovation will face implementation challenges due to one or more gaps in the ecosystem
Innovation Team Potential	0(N), 1(Y)	Whether the team demonstrates potential to successfully execute on the idea it has proposed

4.6 Data

Over 1000 students making up 290 teams submitted ideas through this program over the 2012-2013 time period (Figure 2). A breakup of the submitted ideas by innovation category and the stage to which they reached is presented in Table 4.3.

Table 4.3: Innovation Submissions by Categories

Innovation Category	Sniff Test Stage (Gate A)	Panel Review Stage (Gate 1)	Presentation Round Stage (Gate 2)	Prototype Stage (Gate B)
Mobile Applications	52	17	10	5
Self-service Applications	46	8	4	1
Web Applications	34	7	3	3
Payment Solutions	26	5	5	2
Rural Banking	23	4	3	1
Green Banking	18	7	3	1
Biometric Systems	15	6	3	1
Devices	10	6	0	0
Business Intelligence	6	4	2	0
Next Generation Banking	5	3	1	1
NFC Banking	3	3	2	0
Bluetooth Banking	2	2	1	0
Cloud Computing	2	2	1	1
Branch Banking	1	0	0	0
Internet Banking	1	0	0	0
Others	46	6	1	0
Total	290	80	39	16

Of the 290 submitted ideas, 80 cleared the sniff test stage gate (Gate A) and were evaluated by the review panel (Gate 1); 39 of these ideas cleared Gate 1 and were presented to the Jury (Gate 2) to qualify for prototyping; 16 ideas entered the prototyping stage. Once the prototypes were developed, they were presented to a jury (Gate B) for qualifying for commercialization and funding. To illustrate, in the mobile applications category, 52 ideas were evaluated at the sniff test stage gate (Gate A); 17 of these ideas made it through and were evaluated by the review panel (Gate 1); 10 of these ideas cleared Gate 1 and were presented to the Jury (Gate 2) to qualify for prototyping; 5 ideas entered the prototyping phase and were presented to a jury for qualifying for commercialization and funding (Gate B).

For each of the 80 ideas that were evaluated at Stage Gate 1, the panel gave scores on innovation uniqueness, innovation scalability and innovation business value (Table 4.1). These ratings are based on an analysis of a two page concept note that the teams had submitted as part of team registrations. Text mining and detailed analysis of the comments and texts noted by the review panel and jury provided values for the attributes - Incremental innovation, Regulatory constraint and Ecosystem constraint. In addition, for the 39 ideas that were presented at Gate 2, we meticulously parsed panel feedback on team capabilities to identify high potential teams and categorize them accordingly. These comments were provided by the panel at Stage Gate 2 based on team presentations and the question-answer session following each team′s presentation.

In Table 4.4 we present summary statistics for the 80 ideas that cleared stage gate A and were evaluated in stage gates 1 and 2. A correlation matrix between the attributes is provided in Table 4.5. Finally, to get a sense for how the idea attributes vary by stage gate, we plot them for all the 80 ideas (Figure 4.5).

Table 4.4: Summary Statistics of Innovation Attributes

Innovation Attribute	Variable Name	Mean	Median	Variance	Min.	Max.
Innovation Uniqueness	Uniq	12.31	10	59.79	2	30
Innovation Scalability	Scale	8.62	8	11.24	0	15
Innovation Business Value	Bizval	18.41	20	85.87	2	38
Incremental Innovation	Incrmnt	0.4	0	0.24	0	1
Regulatory Constraints	RegulI	0.08	0	0.07	0	1
Ecosystem Constraints	Ecosys	0.43	0	0.25	0	1
Innovation Team Potential	Teampot	0.29	0	0.21	0	1

Table 4.5: Innovation Attributes Correlation Matrix

Innovation Attribute	Innovation Uniqueness	Innovation Scalability	Business Value	Incremental Innovation	Regulatory Constraint	Ecosystem Constraint	Team Potential
Innovation Uniqueness	1						
Innovation Scalability	0.32	1					
Innovation Business Value	0.39	0.40	1				
Incremental Innovation	-0.43	-0.03	-0.24	1			
Regulatory Constraint	0.11	-0.14	-0.03	-0.23	1		
Ecosystem Constraint	0.09	-0.15	0.06	-0.19	0.04	1	
Innovation Team Potential	0.39	0.33	0.50	-0.29	-0.08	0.07	1

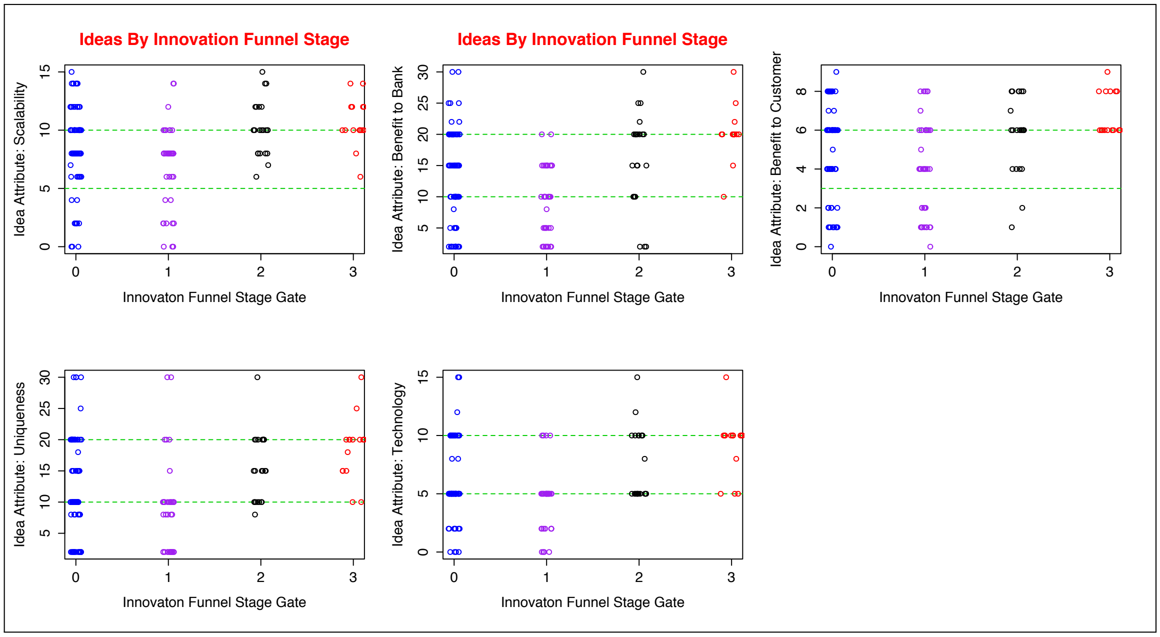


Figure 4.5: Innovation Attributes by Innovation Funnel Stage Gates

4.7 Empirical Model & Estimation

We are modeling the flow of ideas through two stage gates in the innovation funnel. At each stage gate the outcome of interest is whether the idea was selected. The decision makers at both the stage gates are a panel of senior and mid-level managers from the bank. These decisions are discrete sequential choices and can be analyzed empirically.

Consider an innovation i in a two-stage model with a decision made at each stage j to select or reject the idea. Then for a given idea i, i = 1,.... N, the perceived profit at each stage j, j = 1,2 can be expressed as a sum of two components:

$$U_{ij}^{*} = \beta_j X_{ij} + \varepsilon_{ij}$$

Where X_{ij} denotes the observed component which is a known function of various idea attributes, and ε_{ij} is an unobserved random component that captures any unobserved factors affecting the probability of an innovation being selected at the stage gate. Of course, the perceived utility of an innovation is unobservable; however the choice to select or reject an idea is observable. Thus, we can define the binary outcome of selection or rejection as:

$$\mathrm{Y}_{ij} = \begin{cases} 1, & \text{if } U^*_{ij} > 0 \\ 0, & \text{otherwise} \end{cases}$$

In our theoretical development and measures section, we have identified the different factors within the TOE contextual layers that we regard as having a bearing on the utility of each innovation. These factors are summarized in Table 4.2. To model the utility function at each stage gate, we need to identify the innovation attributes that make up the systematic component for each stage gate. In what follows we identify these attributes for each of the stage gates in the innovation funnel. *Stage gate 1 model In the set of explanatory variables for the Stage Gate 1 decision, we include the innovation attributes that were explicitly scored by the review panel - innovation uniqueness($Uniq$), innovation scalability($Scale$) and whether the innovation is incremental($Incrmnt$). Additionally, in line with our theoretical argumentation about the precedence of elimination decision strategies in earlier stages to quickly reduce the consideration set (Figure 4.2), we include regulatory constraints($Regul$) and business value of innovation($Bizval$). Continuing to evaluate innovations that cannot be implemented because they do meet industry or other regulations is a waste of cognitive effort and hence are likely to be eliminated without further thought. Similarly, innovations that do not seem to have a business value upside are likely to be quickly rejected.

On the other hand, evaluating the Ecosystem constraints($Ecosys$) for each innovation is likely an exhaustive process requiring consideration of the various technological and organizational sub-systems within which the innovation must thrive. Thus we contend that this more involved evaluation is akin to an inclusionary decision strategy which in an effort to save cognitive resources is likely to be employed on the reduced set of innovations that reach Stage Gate 2. Thus we express the underlying utility function measuring the propensity of an idea to be selected at Stage Gate 1 as below:

$$U_1^* = \beta_{10} + \beta_{11} * Uniq + \beta_{12} * Scale + \beta_{13} * Incrmnt + \beta_{14} * Bizval + \beta_{15} * Regul + \beta_{16} * Ecosys + \beta_{17} * Teampot + \varepsilon_1$$

Where, ε_1 captures any unobserved factors affecting the probability of an idea being selected in stage gate 1.

The selection outcome depends upon the latent utility being positive as indicated by the outcome variable:

$$Y_1 = \begin{cases} 1, & \text{if } U_1^* > 0 \\ 0, & \text{otherwise} \end{cases}$$

*Stage gate 2 model Stage gate 2 evaluation is very similar to the stage gate 1 evaluation with a panel of senior and mid-level managers reviewing the innovations that were selected at stage gate 1. Hence, similar to stage gate 1, we include the set of innovation attributes, explicitly scored by the review panel - innovation uniqueness($Uniq$), innovation scalability($Scale$) and whether the innovation is incremental($Incrmnt$).

However, there are some important differences between stage gate 1 and 2. First, the number of innovations to be evaluated has been reduced from 80 to 39 and hence more cognitive effort is potentially available per idea. Second, the evaluation is likely to be much more exhaustive since ideas selected at stage gate 2 require commitment of resources from the bank such as the award money for crossing stage gate 1. However, a much more significant resource commitment that the firm offers to the innovations selected at stage gate 2 for the prototyping phase, is managerial time in the form of mentorship for prototype development. The stakes are also higher since the innovations selected after the prototyping phase will be funded by the bank for development into full-fledged commercial startups with the bank being its first customer. In sum, we contend that stage 2 has the characteristics likely to trigger an exhaustive evaluation of ideas. Thus we include Ecosystem constraints($Ecosys$) as one of the covariates.

Another important difference at stage gate 2, as discussed in our hypotheses development, is that the panel in addition to having access to the concept note, also sees the team presentations and cross-examines them with questions. Thus the panel at stage gate 2 has insights into the team's ability in addition to the innovation quality. Thus, for stage gate 2, the team's potential (*Teampot*)is also included as a covariate in the utility function. Thus we express the underlying utility function measuring the propensity of an idea to be selected at Stage Gate 1 as below:

$$U_2^* = \beta_{20} + \beta_{21} * Uniq + \beta_{22} * Scale + \beta_{23} * Incrmnt + \beta_{24} * Ecosys + \beta_{25} * Teampot + \varepsilon_2$$

Where, ε_2, captures any unobserved factors affecting the probability of an idea being selected in stage gate 2.

The selection outcome depends upon the latent utility being positive as indicated by the outcome variable:

$$Y_2 = \begin{cases} 1, & \text{if } U_2^* > 0 \\ 0, & \text{otherwise} \end{cases}$$

Because of the design of the stage gate process, we observe data at stage gate 2 only for ideas that have been selected at stage gate 1. In other words, the innovations observed at stage gate 2 are non-randomly selected from the set of innovations at stage gate 1. To deal with this problem, we assume that the error components are drawn from a bivariate normal distribution, corrected for a sample selection. Additionally, there are reasons to believe that the unobserved components ε_1 and ε_2 are correlated across the decisions at the two stage gates. The decision makers at both stage gates are senior and mid-level managers from the same bank. Both these groups of decision makers have undergone the same orientation and training, and have been acclimatized to the same organizational culture. Thus there is a high chance of correlation in the underlying decision making drivers for these two groups of decision makers. Thus, we assume that the error components are drawn from a bivariate normal distribution, corrected for a sample selection, with a correlation coefficient ρ, such that:

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} = N \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sum \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Thus we jointly estimate the outcomes at both stage gates using a bivariate normal probit model for easy consideration of the correlation between the error terms ε_1 and ε_2. Since there is no closed form solution to this model, a simple analytic method based on minimizing sum of squares is not possible. We use maximum likelihood estimation(MLE) method in which we generate a joint likelihood function for both the stages and maximize it to estimate the coefficients. To specify the joint likelihood function, we start by explicitly specifying possible outcomes for an innovation within the funnel with their associated probabilities (Table 4.6).

Table 4.6: Possible Outcomes For Innovations In Innovation Funnel

Outcome	Y_1	Y_2	Probability	Notation
Innovation Rejected in Stage Gate 1	0	-	$\Pr(Y_1{=}0) = \Pr(U_1^* \leq 0)$	P_0
Selected in Gate 1; Rejected in Gate 2	1	0	$\Pr(Y_1{=}1,Y_2{=}0) = \Pr(U_1^* > 0,\, U_2^* \leq 0)$	P_{10}
Innovation Selected in Stage Gates 1 & 2	1	1	$\Pr(Y_1{=}1,Y_2{=}1) = \Pr(U_1^* > 0,\, U_2^* > 0)$	P_{11}

From the above options and their probabilities, we generate the log-likelihood function as below:

$$\ln L = \sum_{i=1}^{N} Y_{i1} Y_{i2} \ln P_{11} + Y_{i1}(1 - Y_{i2}) \ln P_{10} + (1 - Y_{i1}) \ln P_0$$

This log likelihood function is maximized with respect to the coefficient parameters for both stages and the correlation coefficient ρ to estimate the parameters.

4.8 Results and discussion

Results The results of the MLE estimation are presented in Table 4.7 (Model 1). For checking the robustness of our model, we consider an alternate model (Model 2) in which we abandon our theoretical argumentation for the precedence of elimination strategies in earlier stages. In this model we consider a full model so that business value of an innovation is considered in both stages of the innovation funnel. Additionally, ecosystem constraint is considered as an attribute in gate 1 as well. However, we do not include team potential as a covariate in stage 1 because the review panel at stage 1 does not observe the teams. Further, we do not consider regulatory constraints as a covariate in stage 2 because regulatory constraint is a very strict criteria to avoid. Indeed, out of the 39 innovations that reach stage 2, only one has a regulatory constraint. Hence, due to insufficient variation of this variable, we do not include it as part of our stage 2 model. The regression results for this alternate model consist of a lower likelihood value and a higher AIC and BIC criteria. Additionally, for stage gate 2, we get non-intuitive results such as ecosystem constraints being positively correlated with the probability of idea selection. These findings motivate us to stick with our original model (Model 1) in discussing the results of our estimations.

Table 4.7: Bivariate Probit MLE Estimation Results

		Model 1		Model 2	
Variables	Hyp.(Sign)	Gate1	Gate2	GateI	Gate2
Uniq	H3(+)	0.078* (0.033)	-0.033 (0.052)	0.099** (0.033)	0.001 (0.08)
Scale	H5(+)	0.403*** (0.11)	-0.512*** (0.096)	0.162. (0.089)	0.058 (0.902)
Incrmnt	H4(-)	-1.45** (0.503)	1.721*** (0.035)	-1.077* (0.437)	0.07 (1.625)
Bizval	H6(+)	0.11*** (0.03)	-	0.072** (0.023)	0.057 (0.128)
Regul	H1(-)	-2.407** (0.819)	-	-1.758* (0.747)	-
Ecosys	H2(-)	-	-2.82** (0.934)	-1.127** (0.423)	2.106. (1.26)
Teampot	H7(+)	-	2.084. (1.074)	-	2.426 (2.021)
Constant		-6.137*** (1.322)	6.466*** (0.322)	-3.209** (1.189)	-2.318 (12.784)
Observations		80	39	80	39
LogLik		-35.771		-40.183	
BIC		128.509		146.096	
AIC		97.543		110.366	
ρ		0.873		0.986	

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $\cdot p < 0.1$

Robustness check A close look at Table 4.2 reveals that the measures for innovation uniqueness, scalability and business value are weighted differently. For instance Innovation uniqueness ranges from 0-30 while Innovation scalability ranges from 0-15. In order to draw meaningful conclusions about the relative influence of each of these variables, we will need to bring these three variables in proportion with one another. Thus, we standardize each of these three variables to a value between 0-1 and regress our models again. The estimated coefficients, their values signs and significance levels are close to those presented in Table 4.7. Thus we can conclude that our results are robust between the use of standardized and non-standardized variables.

Discussion The findings support several of our hypothesis while throwing up some surprises. The first significant result is a validation of our propositions about cognitively less costly elimination strategies being used in earlier stages of innovation funnels to reduce the consideration followed by employing cognitively more costly choice strategies in the later stages. Thus, regulatory constraints which are a cognitively less costly decision heuristic seem to be employed in gate 1. On the other hand, ecosystem constraints which are cognitively more costly seem to be employed as a decision heuristic for eliminating ideas in stage gate 2. Even more interestingly, none of these elimination criteria were explicitly provided to the review panel for selecting or rejecting ideas. The criteria provided to them was a rubric consisting of idea characteristics such as its uniqueness and scalability. Indeed, the panel seems to have consciously or subconsciously *improvised* to include elimination criteria in their decision heuristics to quickly reduce the consideration set in stage gate 1. This preference for elimination strategies in earlier stages is in line with findings in studies analyzing the relative importance of exclusion versus inclusion choice strategies as discussed in our theory development.

A second surprising finding is a stage gate effect in the reversal of preference for unique versus incremental innovations. Thus we find that unique ideas are preferred in Gate 1 but not in Gate 2. Similarly, ideas with higher scalability which arguably correlates with more ambitious ideas are favored in the gate 1 but not in gate 2. On the other hand, incremental ideas seem likely to be rejected in gate 1 but are more likely to be selected in gate 2. All these finding seem to suggest that firms in the initial stages wish to be innovative and end up choosing ambitious and innovative ideas. However, when the psychological distance from the point of committing resources is removed, it gets cold feet and reverts to the safe grounds of familiar and incremental innovations. Thus there seems to be a preference for the innovative in the early stages and a presence for the conservative in stages closer to resource commitment. We believe that this to our knowledge is a first empirical evidence to demonstrate a stage gate effect in selection of innovations. It also seems to propose a potential candidate mechanism for how established innovative firms might develop game changing ideas with a goal to be innovative but consistently abandon it when the pivotal point of committing resources is near. In what follows, we examine each of the explanatory variables and their correlation with probability of idea selection.

Idea Uniqueness. Idea uniqueness is positively correlated with the probability of idea selection in stage gate 1 at the 10% significance level. However, the correlation is not statistically significant for stage gate 2. Thus H2 is supported for stage gate 1 but not for stage gate 2. This might be explained by the desire of the decision makers at the earlier stages to align with the mandate of open innovation to promote novelty and uniqueness. However, in the later stages which are psychologically and temporarily very close to commitment of resources, the uniqueness of an idea might seem much less salient than aspects which speak to the ability of the idea to be executed successfully. In the later stages the focus of the evaluation thus seems to shift from the why to the how.

Idea scalability. Idea Scalability is positively correlated with the probability of idea selection in stage gate 1 at the 1% significance level. However in stage 2 it is negatively and significantly correlated with the probability of idea selection. Thus Hypothesis H7 is supported for stage gate 1. However, for stage gate 2, the signs are surprisingly reversed.Though the negative sign on the coefficient for scalability for stage gate 2 is puzzling, further reflection tends to provide possible reasoning that might lead to such a bias. In the later stages the decision to choose an idea will result in the bank committing significant resources to a certain path. Though in the earlier stages, the decision makers align with the desirable goal of pursuing the most scalable ideas, when it comes time to commit resources, the decision makers are justifiably apprehensive of committing to highly scalable ideas. This line of reasoning seems increasingly justifiable when we take into consideration the fact that the team leading these ventures are not seasoned entrepreneurs but engineering undergraduates who are still in college.

Regulatory constraints. Regulatory constraints are negatively and significantly correlated with the probability of idea selection for both stage gates 1 and 2. This supports hypothesis H1a. As discussed in theory and hypothesis building, regulatory constraints are an instance of an exclusion criteria that are firm and impermeable. Thus, they are likely to be used as a heuristic to quickly and costlessly eliminate ideas from the consideration set. The implication of this finding is that similar mental short-cuts (unrelated to regulatory constraints) if frequently employed by decision makers might explain disruptive innovations being rejected from the innovation funnel prematurely.

Ecosystem constraints. Ecosystem gaps are negatively correlated with the probability of idea selection in stage gate 2 at the 5% significance level. However, the correlation is not statistically significant for stage gate 1. Thus H1b is supported for stage gate 2 but not for stage gate 1. Ecosystem gaps may not have an influence on idea selection in stage gate 1 because of two possible reasons. First, senior managers because of their experience and vantage point are better able to spot gaps in the ecosystem that might becoming stumbling blocks to the implementation of an idea. Alternatively, this might be more of a ′stage effect′ in that the ecosystem gaps becomes more apparent only after the details of the idea have been more fully fleshed out in the later stages of the innovation funnel.

Incremental innovation. An innovation′s sustaining nature is negatively and significantly correlated with the probability of idea selection in stage gate 1. However in stage 2 it is positively and significantly correlated with the probability of idea selection. Thus Hypothesis H4 is supported for stage gate 2. However, for stage gate 1, there is a surprising reversal of signs. This might be explained by the psychological distance to commitment in the earlier stages of the funnel. When commitment to a certain path of action is in the distant future, the decision makers prefer the desirable goal of promoting novel ideas by weeding sustaining innovations. However, in the later stages, when faced with a decision which will commit the organization to a certain path of action, the decision makers prefer the safer route of pursuing innovations they are familiar with. This difference in approach in the two stages can be explained as a stage gate effect in that as the commitment implications of a decision become more near, decision makers tend to be more conservative. Alternatively, this might be explained by the claim that senior managers are more conservative than younger managers due to a tenure effect.

Innovation business value. A surprising finding is that neither the benefit to the bank or benefit to the customer were statistical significant for idea selection. Thus both H5a and H5b are not supported. Though at first glance this seems counterintuitive, a deeper analysis reveals that these two criteria are cognitively very costly since they require predictions of possible future high stakes scenarios. Further, these predictions are based on evaluations of multiple considerations related to customers, business, market and industry. Hence it may not be surprising if the decision makers seek refuge in cognitively less costly attributes such as those which require a simple check for the presence or absence of an attribute like regulatory constraints. It is important to note that since we did not do a process tracing analysis of decision making, we cannot lay claim to explaining the mechanism by which the decisions were made. However, previous process tracing studies have found support for the primacy and precedence of elimination strategies over selection strategies which affords it as a candidate explanation for this surprising finding.

Innovation team potential. Team execution potential is positively and significantly correlated with the probability of idea selection at stage gate 2. This supports hypothesis H8 and is consistent with our expectations. The decision makers are not only choosing the idea but also the team that will execute it and thus teams that show the ability to execute on their plans will have a higher probability of being selected.

4.9 Conclusion

With the dawn of the information age and digitization, software is eating the world. This implies that knowledge and thus innovation is increasingly encapsulated in digital rather than physical artifacts. Since the digital medium is relatively free of the barriers of scale and independent of domain, it opens up the possibility of democratization of innovation. However, the promise of open innovation is contingent on the firm picking the right ideas from those pouring in from external sources. This explains our departure from the traditional focus in extant research on the ideation process. In this study, we set out to study a stage gate model of an innovation funnel to investigate the decision making criteria for picking technological innovations in an open innovation funnel. We specifically try to identify the attributes of an idea that are associated with its selection. We carry out this investigation in a framework created by synthesizing theories from behavioral and institutional economics. To our knowledge, these are first such attempts in the IS innovation literature and we present our findings as a call to further investigations in the technological innovation selection process in open innovation funnels.

Overall, we find support for the presence of a combination of elimination and compensatory selection strategies. Additionally, the elimination strategies seem to have primacy. This is in agreement with decision research which has found that as the number of alternatives and the number of evaluation attributes increase, elimination strategies become increasingly attractive due to their low cognitive cost. Thus the common practice of designing a multi-attribute criteria set to be used for selecting ideas seems contrary to how people choose. The managerial implication of this finding is that when designing criteria for selecting innovation in innovation funnels,firms, should pay equal if not more emphasis on identifying elimination criteria in addition to identifying election criteria. This we believe will have two advantages to the firm. First, ideas that do not meet the minimum requirements, will be eliminated right away thus not taking up scarce management attention till the very end. Secondly, defining a set of elimination criteria will force the firm to articulate where it lies on the continuum between innovativeness and feasibility. This would avoid the counterproductive tendency of paying lip service to innovation while staying stuck in a status quo mindset.

Secondly, we found support for a shift in the evaluation criteria between early and late stage gates of the innovation funnels. In the earlier stages, there is a tendency for the decision makers to side with the desirable goal of being innovative and radical. However, in the later stage gates, we find support for status quo centric attributes such as feasibility, familiarity, and team′s ability to execute, gaining salience. In line with this innovativeness versus feasibility dichotomy, we find that sustaining innovations are favored in later stages while they are rejected in earlier stages. Similarly, highly scaleable ideas are favored in earlier stages, but when closer to implementation in later stages, the decision makers turn conservative towards highly scalable ideas. Additionally, unique ideas are favored in the early stages but not in the later stages.

In sum, these findings suggest that even if innovative ideas might enter the open innovation funnel, an initial enthusiasm for innovative ideas might be tempered in later stages by a concern for implementation and feasibility. The managerial implication of these findings is that these tendencies towards the status quo in later stages might explain why established firms end up ignoring disruptive innovations. Further, being aware of and accepting this status quo bias could point firms to potential solutions and strategies. For one, the firm could deliberately create space for radical ideas by having two categories of selected ideas. Those that will be implemented right away and thus need to have the feasibility criteria. And another category of ideas which do not have the feasibility constraint and are thus evaluated only on their innovativeness. The firm could possibly have a skunk work where these innovative ideas are field tested.

4.10 Contributions and limitations

We believe that this study makes several contributions to the IS innovation literature. First, to our knowledge, this study is a first in the IS innovation literature to shift the focus from idea generation to idea selection in an innovation funnel. Second, to our knowledge, it is one of few IS innovation papers which bring to bear behavioral economics theories in a significant way to explain decision making within an open innovation funnel. Third, this study to our knowledge is a first in the IS innovation literature to draw out the differences in decision making criteria between early and later stages of the innovation funnel. More specifically, this study highlights the division of labor between exclusionary and inclusionary decision criteria in selecting ideas. Most significantly, this is a first study in the IS innovation literature to find empirical support for a stage effect in the propensity to be innovative, thus directing attention to the later stages of an innovation funnel as the most likely weak link where the ball of innovative ideas are dropped.

A limitation of our study is that it concerns a single firm in an emerging economy, thus limiting the generalizability of our findings. However, the firm is one of the largest private banks in India with global operations and widely recognized for being a technology pioneer. This does make it a close to ideal case to study open innovation in incumbent firms. Another factor in our favor is that the innovations are funded and commercialized by the bank. In this sense, unlike some studies that examine 'toy contests' our setting is possibly a close enactment of the innovation selection process in firms. Secondly, this innovation contest is focused on information technology centric innovations. Thus, it is not clear whether our findings translate to non-technology sectors such as drug discovery. Finally, further scrutiny might be needed before our findings can be applied to similar processes in which a subset of opportunities are selected from a wider set such as technology acquisitions, startup funding and venture capital deals.

BIBLIOGRAPHY

BIBLIOGRAPHY

[Abbasi et al., 2012] Abbasi, A., Albrecht, C., Vance, A., and Hansen, J. (2012). Metafraud: a meta-learning framework for detecting financial fraud. *Mis Quarterly*, 36(4):1293–1327.

[Abbey and Dickson, 1983] Abbey, A. and Dickson, J. W. (1983). R&d work climate and innovation in semiconductors. *Academy of Management Journal*, 26(2):362–368.

[Adner, 2006] Adner, R. (2006). Match your innovation strategy to your innovation ecosystem. *Harvard business review*, 84(4):98.

[Adner, 2012] Adner, R. (2012). *The wide lens: A new strategy for innovation*. Penguin UK.

[Adner and Kapoor, 2010] Adner, R. and Kapoor, R. (2010). Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations. *Strategic management journal*, 31(3):306–333.

[Adomavicius et al., 2009] Adomavicius, G., Gupta, A., and Zhdanov, D. (2009). Designing intelligent software agents for auctions with limited information feedback. *Information Systems Research*, 20(4):507–526.

[Agarwal et al., 2014] Agarwal, R., Bapna, R., Yong Goh, K., Ghose, A., Shmueli, G., and Slaughter, S. (2014). Does growing demand for data science create new opportunities for information systems? In *Thirty Fifth International Conference on Information Systems*. Thirty Fifth International Conference on Information Systems.

[Agarwal and Prasad, 1997] Agarwal, R. and Prasad, J. (1997). The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision sciences*, 28(3):557–582.

[Ajzen and Madden, 1986] Ajzen, I. and Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of experimental social psychology*, 22(5):453–474.

[Alavi and Leidner, 2001] Alavi, M. and Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*, pages 107–136.

[Albert et al., 2004] Albert, T. C., Goes, P. B., and Gupta, A. (2004). Gist: a model for design and management of content and interactivity of customer-centric web sites. *MIS Quarterly*, pages 161–182.

[Amabile, 1996] Amabile, T. M. (1996). *Creativity in context: Update to" the social psychology of creativity."*. Westview press.

[Amit and Schoemaker, 1993] Amit, R. and Schoemaker, P. J. (1993). Strategic assets and organizational rent. *Strategic management journal*, 14(1):33–46.

[Anderson, 1982] Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, 89(4):369.

[Anderson, 1996] Anderson, J. R. (1996). Act: A simple theory of complex cognition. *American Psychologist*, 51(4):355.

[Ang et al., 2011] Ang, S., Banker, R., Bapna, R., Slaughter, S., and Wattal, S. (2011). Human capital of it professionals: A research agenda.

[Ang and Low, 2000] Ang, S. H. and Low, S. Y. (2000). Exploring the dimensions of ad creativity. *Psychology & Marketing*, 17(10):835–854.

[Arthur, 1994] Arthur, W. B. (1994). *Increasing returns and path dependence in the economy*. University of Michigan Press.

[Arthur, 2009] Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves*. Simon and Schuster.

[Ashby et al., 1956] Ashby, W. R. et al. (1956). *An introduction to cybernetics*, volume 2. Chapman & Hall London.

[Bailetti and Callahan, 1992] Bailetti, A. J. and Callahan, J. R. (1992). Assessing the impact of university interactions on an r&d organization. *R&D Management*, 22(2):145–156.

[Balachandra, 1984] Balachandra, R. (1984). Critical signals for making go/nogo decisions in new product development. *Journal of Product Innovation Management*, 1(2):92–100.

[Bapna et al., 2013] Bapna, R., Langer, N., Mehra, A., Gopal, R., and Gupta, A. (2013). Human capital investments and employee performance: an analysis of it services industry. *Management Science*, 59(3):641–658.

[Barlow, 2000] Barlow, C. M. (2000). Deliberate insight in team creativity. *The Journal of Creative Behavior*, 34(2):101–117.

[Baskerville and Wood-Harper, 1998] Baskerville, R. and Wood-Harper, A. T. (1998). Diversity in information systems action research methods. *European Journal of Information Systems*, 7(2):90–107.

[Basu and Blanning, 2003] Basu, A. and Blanning, R. W. (2003). Synthesis and decomposition of processes in organizations. *Information Systems Research*, 14(4):337–355.

[Baumann and Martignoni, 2011] Baumann, O. and Martignoni, D. (2011). Evaluating the new: The contingent value of a pro-innovation bias. *Schmalenbach Business Review (sbr)*, 63(4).

[Besemer and O'QUIN, 1986] Besemer, S. and O'QUIN, K. (1986). Analyzing creative products: Refinement and test of a judging instrument. *The Journal of Creative Behavior*, 20(2):115–126.

[Billings and Marcus, 1983] Billings, R. S. and Marcus, S. A. (1983). Measures of compensatory and noncompensatory models of decision behavior: Process tracing versus policy capturing. *Organizational Behavior and Human Performance*, 31(3):331–352.

[Bitterer, 2011] Bitterer, A. (2011). Hype cycle for business intelligence. *Gartner, Inc., Stamford, CT.*

[Blair and Mumford, 2007] Blair, C. S. and Mumford, M. D. (2007). Errors in idea evaluation: Preference for the unoriginal? *The journal of creative behavior*, 41(3):197–222.

[Booz, 1982] Booz (1982). *New products management for the 1980s.* Booz, Allen & Hamilton.

[Boulding, 1996] Boulding, K. E. (1996). 16 the economics of the coming spaceship earth. *Valuing the earth: economics, ecology, ethics*, page 297.

[Bresnahan et al., 1999] Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. (1999). Information technology, workplace organization and the demand for skilled labor: firm-level evidence. Technical report, National bureau of economic research.

[Brynjolfsson and McAfee, 2011] Brynjolfsson, E. and McAfee, A. (2011). Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy.

[Brynjolfsson and McAfee, 2014] Brynjolfsson, E. and McAfee, A. (2014). *The second machine age: work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

[Burstein et al., 1999] Burstein, F., Gregor, S., et al. (1999). The systems development or engineering approach to research in information systems: An action research perspective. In *Proceedings of the 10th Australasian Conference on Information Systems*, pages 122–134. Victoria University of Wellington, New Zealand.

[Cady and Valentine, 1999] Cady, S. H. and Valentine, J. (1999). Team innovation and perceptions of consideration what difference does diversity make? *Small group research*, 30(6):730–750.

[Chau and Xu, 2012] Chau, M. and Xu, J. (2012). Business intelligence in blogs: Understanding consumer interactions and communities. *MIS quarterly*, 36(4):1189–1216.

[Chaudhuri et al., 2011] Chaudhuri, S., Dayal, U., and Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, 54(8):88–98.

[Chen et al., 2012] Chen, H., Chiang, R. H., and Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 36(4):1165–1188.

[Chesbrough and Brunswicker, 2014] Chesbrough, H. and Brunswicker, S. (2014). A fad or a phenomenon?: The adoption of open innovation practices in large firms. *Research-Technology Management*, 57(2):16–25.

[Chesbrough et al., 2006] Chesbrough, H., Vanhaverbeke, W., and West, J. (2006). *Open innovation: Researching a new paradigm*. Oxford university press.

[Chesbrough, 2003] Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.

[Chiang and Mookerjee, 2004] Chiang, I. R. and Mookerjee, V. S. (2004). A fault threshold policy to manage software development projects. *Information Systems Research*, 15(1):3–21.

[Chiesa and Manzini, 1998] Chiesa, V. and Manzini, R. (1998). Organizing for technological collaborations: a managerial perspective. *R&D Management*, 28(3):199–212.

[Christensen, 1997] Christensen, C. M. (1997). The innovator's dilemmaharvard business school press. *Boston, MA*.

[Christiaans, 2002] Christiaans, H. H. (2002). Creativity as a design criterion. *Communication Research Journal*, 14(1):41–54.

[Cohen and Levinthal, 1990] Cohen, W. M. and Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative science quarterly*, pages 128–152.

[Conway, 1995] Conway, S. (1995). Informal boundary-spanning communication in the innovation process: An empirical study. *Technology Analysis & Strategic Management*, 7(3):327–342.

[Cooper, 2008] Cooper, R. G. (2008). Perspective: The stage-gate® idea-to-launch process—update, what's new, and nexgen systems*. *Journal of Product Innovation Management*, 25(3):213–232.

[Cooper and Kleinschmidt, 1990] Cooper, R. G. and Kleinschmidt, E. J. (1990). New product success factors: a comparison of 'kills' versus successes and failures. *R&D Management*, 20(1):47–63.

[Cooper and Kleinschmidt, 1993] Cooper, R. G. and Kleinschmidt, E. J. (1993). Stage gate systems for new product success. *Marketing Management*, 1(4):20–29.

[Craig et al., 2012] Craig, E., Smith, D., Mulani, N. P., and Thomas, R. J. (2012). Where will you find your analytics talent? *Outlook*.

[Daft and Weick, 1984] Daft, R. L. and Weick, K. E. (1984). Toward a model of organizations as interpretation systems. *Academy of management review*, 9(2):284–295.

[Danneels, 1998] Danneels, E. (1998). A competence-based new product typology. *ISBM Report 17-1998*.

[Davenport, 2006] Davenport, T. H. (2006). Competing on analytics. *harvard business review*, 84(1):98.

[Davenport, 2013] Davenport, T. H. (2013). *Process innovation: reengineering work through information technology*. Harvard Business Press.

[Davenport et al., 2013] Davenport, T. H., Barth, P., and Bean, R. (2013). How big datais different.

[Davenport and Patil, 2012] Davenport, T. H. and Patil, D. (2012). Data scientist. *Harvard Business Review*, 90:70–76.

[Davis, 1989] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pages 319–340.

[Dean et al., 2006] Dean, D. L., Hender, J. M., Rodgers, T. L., and Santanen, E. L. (2006). Identifying quality, novel, and creative ideas: Constructs and scales for idea evaluation. *Journal of the Association for Information Systems*, 7(10).

[Dedehayir and Mäkinen, 2011] Dedehayir, O. and Mäkinen, S. J. (2011). Determining reverse salient types and evolutionary dynamics of technology systems with performance disparities. *Technology Analysis & Strategic Management*, 23(10):1095–1114.

[Dewar and Dutton, 1986] Dewar, R. D. and Dutton, J. E. (1986). The adoption of radical and incremental innovations: an empirical analysis. *Management science*, 32(11):1422–1433.

[Diehl and Stroebe, 1987] Diehl, M. and Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of personality and social psychology*, 53(3):497.

[DiMaggio and Powell, 1983] DiMaggio, P. J. and Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review*, pages 147–160.

[Ding and Eliashberg, 2002] Ding, M. and Eliashberg, J. (2002). Structuring the new product development pipeline. *Management Science*, 48(3):343–363.

[Ericsson et al., 1993] Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3):363.

[Ethiraj et al., 2005] Ethiraj, S. K., Kale, P., Krishnan, M. S., and Singh, J. V. (2005). Where do capabilities come from and how do they matter? a study in the software services industry. *Strategic Management Journal*, 26(1):25–45.

[Ettlie and Rubenstein, 1987] Ettlie, J. E. and Rubenstein, A. H. (1987). Firm size and product innovation. *Journal of Product Innovation Management*, 4(2):89–108.

[Finke et al., 1992] Finke, R. A., Ward, T. B., and Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. MIT press Cambridge, MA.

[Ford et al., 1989] Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., and Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes*, 43(1):75–117.

[Foster, 1986] Foster, R. N. (1986). *Innovation: The attacker's advantage*, volume 152. Summit Books New York.

[Franke and Hienerth, 2006] Franke, N. and Hienerth, C. (2006). Prädikatoren der qualität von geschäftsideen: eine empirische analyse eines online-ideen-forums. *Zeitschrift für Betriebswirtschaft, Special Issue*, 76(4):47–68.

[Gallagher et al.,] Gallagher, K. P., Kaiser, K. M., Simon, J. C., Beath, C. M., and Goles, T. The requisite variety of skills for it professionals.

[Gallupe et al., 1992] Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., and Nunamaker, J. F. (1992). Electronic brainstorming and group size. *Academy of Management Journal*, 35(2):350–369.

[Garud et al., 1997] Garud, R., Nayyar, P. R., and Shapira, Z. B. (1997). *Technological innovation: Oversights and foresights*. Cambridge University Press.

[Gatignon et al., 2002] Gatignon, H., Tushman, M. L., Smith, W., and Anderson, P. (2002). A structural approach to assessing innovation: Construct development of innovation locus, type, and characteristics. *Management Science*, 48(9):1103–1122.

[Ghemawat, 1991] Ghemawat, P. (1991). *Commitment*. Simon and Schuster.

[Gill, 1995] Gill, T. G. (1995). Early expert systems: Where are they now? *MIS quarterly*, pages 51–81.

[Girotra et al., 2010] Girotra, K., Terwiesch, C., and Ulrich, K. T. (2010). Idea generation and the quality of the best idea. *Management Science*, 56(4):591–605.

[Glasmeier, 1991] Glasmeier, A. (1991). Technological discontinuities and flexible production networks: The case of switzerland and the world watch industry. *Research policy*, 20(5):469–485.

[Gopal et al., 2003] Gopal, A., Sivaramakrishnan, K., Krishnan, M. S., and Mukhopadhyay, T. (2003). Contracts in offshore software development: An empirical analysis. *Management Science*, 49(12):1671–1683.

[Green et al., 1995] Green, S. G., Gavin, M. B., and Aiman-Smith, L. (1995). Assessing a multidimensional measure of radical technological innovation. *Engineering Management, IEEE Transactions on*, 42(3):203–214.

[Greenwood and Hinings, 1996] Greenwood, R. and Hinings, C. R. (1996). Understanding radical organizational change: Bringing together the old and the new institutionalism. *Academy of management review*, 21(4):1022–1054.

[Gregor and Benbasat, 1999] Gregor, S. and Benbasat, I. (1999). Explanations from intelligent systems: Theoretical foundations and implications for practice. *MIS quarterly*, pages 497–530.

[Gregory et al., 1991] Gregory, R., Kunreuther, H., Easterling, D., and Richards, K. (1991). Incentives policies to site hazardous waste facilities. *Risk Analysis*, 11(4):667–675.

[Griffiths-Hemans and Grover, 2006] Griffiths-Hemans, J. and Grover, R. (2006). Setting the stage for creative new products: investigating the idea fruition process. *Journal of the Academy of Marketing Science*, 34(1):27–39.

[Hagedoorn, 1993] Hagedoorn, J. (1993). Understanding the rationale of strategic technology partnering: Nterorganizational modes of cooperation and sectoral differences. *Strategic management journal*, 14(5):371–385.

[Haour, 1992] Haour, G. (1992). Stretching the knowledge-base of the enterprise through contract research. *R&D Management*, 22(2):177–182.

[Harter et al., 2000] Harter, D. E., Krishnan, M. S., and Slaughter, S. A. (2000). Effects of process maturity on quality, cycle time, and effort in software product development. *Management Science*, 46(4):451–466.

[Hathaway, 2001] Hathaway, O. A. (2001). Path dependence in the law: The course and pattern of legal change in a common law system. *As published in Iowa Law Review*, 86.

[Heller et al., 2002] Heller, D., Levin, I. P., and Goransson, M. (2002). Selection of strategies for narrowing choice options: Antecedents and consequences. *Organizational Behavior and Human Decision Processes*, 89(2):1194–1213.

[Henderson, 1993] Henderson, R. (1993). Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry. *The RAND Journal of Economics*, pages 248–270.

[Henderson and Clark, 1990] Henderson, R. M. and Clark, K. B. (1990). Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative science quarterly*, pages 9–30.

[Hogarth, 1987] Hogarth, R. M. (1987). *Judgement and choice: The psychology of decision .* John Wiley & Sons.

[Hu et al., 2012] Hu, D., Zhao, J. L., Hua, Z., and Wong, M. C. (2012). Network-based modeling and analysis of systemic risk in banking systems. *MIS Quarterly*, 36(4):1269–1291.

[Hughes, 1993] Hughes, T. P. (1993). *Networks of power: electrification in Western society, 1880-1930*. JHU Press.

[Iansiti, 1997] Iansiti, M. (1997). From technological potential to product performance: an empirical analysis. *Research policy*, 26(3):345–365.

[Ingham and Mothe, 1998] Ingham, M. and Mothe, C. (1998). How to learn in r&d partnerships? *R&D Management*, 28(4):249–261.

[Johne and Snelson, 1988] Johne, F. A. and Snelson, P. A. (1988). Success factors in product innovation: a selective review of the literature. *Journal of product innovation management*, 5(2):114–128.

[Kogut and Kulatilaka, 2001] Kogut, B. and Kulatilaka, N. (2001). Capabilities as real options. *Organization Science*, 12(6):744–758.

[Kornish and Ulrich, 2011] Kornish, L. J. and Ulrich, K. T. (2011). Opportunity spaces in innovation: Empirical analysis of large samples of ideas. *Management Science*, 57(1):107–128.

[Krishnan and Kellner, 1999] Krishnan, M. S. and Kellner, M. I. (1999). Measuring process consistency: Implications for reducing software defects. *Software Engineering, IEEE Transactions on*, 25(6):800–815.

[Krishnan et al., 2000] Krishnan, M. S., Kriebel, C. H., Kekre, S., and Mukhopadhyay, T. (2000). An empirical analysis of productivity and quality in software products. *Management science*, 46(6):745–759.

[Krishnan and Ulrich, 2001] Krishnan, V. and Ulrich, K. T. (2001). Product development decisions: A review of the literature. *Management science*, 47(1):1–21.

[Kristensson et al., 2004] Kristensson, P., Gustafsson, A., and Archer, T. (2004). Harnessing the creative potential among users*. *Journal of product innovation management*, 21(1):4–14.

[Kuan and Chau, 2001] Kuan, K. K. and Chau, P. Y. (2001). A perception-based model for edi adoption in small businesses using a technology–organization–environment framework. *Information & Management*, 38(8):507–521.

[Kugel, 1988] Kugel, P. (1988). How long will computers stay" dumb"? *MIS Quarterly*, pages 1–3.

[Kuhn, 1962] Kuhn, T. S. (1962). The structure of scientific revolutions.

[Langlois, 1997] Langlois, R. N. (1997). Cognition and capabilities: opportunities seized and missed in the history of the computer industry. *Technological innovation*, pages 71–94.

[Langlois, 2003] Langlois, R. N. (2003). The vanishing hand: the changing dynamics of industrial capitalism. *Industrial and corporate change*, 12(2):351–385.

[LaValle et al., 2013] LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., and Kruschwitz, N. (2013). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, 21.

[Leidner and Kayworth, 2006] Leidner, D. E. and Kayworth, T. (2006). Review: a review of culture in information systems research: toward a theory of information technology culture conflict. *MIS quarterly*, 30(2):357–399.

[Leonard-Barton, 1988] Leonard-Barton, D. (1988). Implementation as mutual adaptation of technology and organization. *Research policy*, 17(5):251–267.

[Levinthal and March, 1993] Levinthal, D. A. and March, J. G. (1993). The myopia of learning. *Strategic management journal*, 14(S2):95–112.

[Lilien et al., 2002] Lilien, G. L., Morrison, P. D., Searls, K., Sonnack, M., and Hippel, E. v. (2002). Performance assessment of the lead user idea-generation process for new product development. *Management science*, 48(8):1042–1059.

[Lim et al., 2011] Lim, J.-H., Stratopoulos, T. C., and Wirjanto, T. S. (2011). Path dependence of dynamic information technology capability: An empirical investigation. *Journal of Management Information Systems*, 28(3):45–84.

[Lucas Jr and Goh, 2009] Lucas Jr, H. C. and Goh, J. M. (2009). Disruptive technology: How kodak missed the digital photography revolution. *The Journal of Strategic Information Systems*, 18(1):46–55.

[MacCrimmon and Wagner, 1994] MacCrimmon, K. R. and Wagner, C. (1994). Stimulating ideas through creative software. *Management science*, 40(11):1514–1532.

[MacMillan et al., 1986] MacMillan, I. C., Siegel, R., and Narasimha, P. (1986). Criteria used by venture capitalists to evaluate new venture proposals. *Journal of Business venturing*, 1(1):119–128.

[MacMillan et al., 1987] MacMillan, I. C., Zemann, L., and Subbanarasimha, P. (1987). Criteria distinguishing successful from unsuccessful ventures in the venture screening process. *Journal of business venturing*, 2(2):123–137.

[Malhotra, 1982] Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of consumer research*, pages 419–430.

[Mayer, 1999] Mayer, R. E. (1999). 22 fifty years of creativity research. *Handbook of creativity*, page 449.

[McAfee et al., 2012] McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., and Barton, D. (2012). Big data. *The management revolution. Harvard Bus Rev*, 90(10):61–67.

[Mehra et al., 2014] Mehra, A., Langer, N., Bapna, R., and Gopal, R. D. (2014). Estimating returns to training in the knowledge economy: A firm level analysis of small and medium enterprises. *Mis Quarterly*, 38(3).

[Melville et al., 2004] Melville, N., Kraemer, K., and Gurbaxani, V. (2004). Review: Information technology and organizational performance: An integrative model of it business value. *MIS quarterly*, 28(2):283–322.

[Meyer et al., 2014] Meyer, G., Adomavicius, G., Johnson, P. E., Elidrisi, M., Rush, W. A., Sperl-Hillen, J. M., and O'Connor, P. J. (2014). A machine learning approach to improving dynamic decision making. *Information Systems Research*, 25(2):239–263.

[Miles et al., 1978] Miles, R. E., Snow, C. C., Meyer, A. D., and Coleman, H. J. (1978). Organizational strategy, structure, and process. *Academy of management review*, 3(3):546–562.

[Mithas and Krishnan, 2008] Mithas, S. and Krishnan, M. S. (2008). Human capital and institutional effects in the compensation of information technology professionals in the united states. *Management Science*, 54(3):415–428.

[More, 1982] More, R. A. (1982). Risk factors in accepted and rejected new industrial products. *Industrial Marketing Management*, 11(1):9–15.

[Morrison and George, 1995] Morrison, J. and George, J. F. (1995). Exploring the software engineering component in mis research. *Communications of the ACM*, 38(7):80–91.

[Nagasundaram and Bostrom, 1994] Nagasundaram, M. and Bostrom, R. P. (1994). The structuring of creative processes using gss: a framework for research. *Journal of Management Information Systems*, 11(3):87–114.

[Narula and Hagedoorn, 1999] Narula, R. and Hagedoorn, J. (1999). Innovating through strategic alliances: moving towards international partnerships and contractual agreements. *Technovation*, 19(5):283–294.

[Nelson, 1990] Nelson, D. L. (1990). Individual adjustment to information-driven technologies: a critical review. *MIS quarterly*, pages 79–98.

[Niu and Sternberg, 2001] Niu, W. and Sternberg, R. J. (2001). Cultural influences on artistic creativity and its evaluation. *International journal of psychology*, 36(4):225–241.

[Nord and Tucker, 1987] Nord, W. R. and Tucker, S. (1987). *Implementing routine and radical innovations*. Lexington Books Lexington, MA.

[Nunamaker Jr and Chen, 1990] Nunamaker Jr, J. F. and Chen, M. (1990). Systems development in information systems research. In *System Sciences, 1990., Proceedings of the Twenty-Third Annual Hawaii International Conference on*, volume 3, pages 631–640. IEEE.

[Oliver, 1991] Oliver, C. (1991). Strategic responses to institutional processes. *Academy of management review*, 16(1):145–179.

[Orlikowski and Iacono, 2001] Orlikowski, W. J. and Iacono, C. S. (2001). Research commentary: Desperately seeking the it in it researcha call to theorizing the it artifact. *Information systems research*, 12(2):121–134.

[Osborn, 1953] Osborn, A. F. (1953). Applied imagination.

[Page, 1999] Page, S. E. (1999). On the emergence of cities. *Journal of Urban Economics*, 45(1):184–208.

[Park et al., 2012] Park, S.-H., Huh, S.-Y., Oh, W., and Han, S. P. (2012). A social network-based inference model for validating customer profile data. *MIS Quarterly*, 36(4):1217–1237.

[Payne, 1982] Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin*, 92(2):382.

[Payne et al., 1993] Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.

[Piccoli and Ives, 2005] Piccoli, G. and Ives, B. (2005). Review: It-dependent strategic initiatives and sustained competitive advantage: a review and synthesis of the literature. *Mis Quarterly*, 29(4):747–776.

[Piller and Walcher, 2006] Piller, F. T. and Walcher, D. (2006). Toolkits for idea competitions: a novel method to integrate users in new product development. *r&D management*, 36(3):307–318.

[Plucker et al., 2004] Plucker, J. A., Beghetto, R. A., and Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? potentials, pitfalls, and future directions in creativity research. *Educational Psychologist*, 39(2):83–96.

[Purao et al., 2003] Purao, S., Storey, V. C., and Han, T. (2003). Improving analysis pattern reuse in conceptual design: Augmenting automated processes with supervised learning. *Information Systems Research*, 14(3):269–290.

[Ramasubbu et al., 2008] Ramasubbu, N., Mithas, S., and Krishnan, M. S. (2008). High tech, high touch: The effect of employee skills and customer heterogeneity on customer satisfaction with enterprise system support services. *Decision Support Systems*, 44(2):509–523.

[Reitzig, 2011] Reitzig, M. (2011). ls your company choosing the best lnnovation ideas? *MIT Sloan Management Review*.

[Rietzschel et al., 2010] Rietzschel, E. F., Nijstad, B. A., and Stroebe, W. (2010). The selection of creative ideas after individual idea generation: Choosing between creativity and impact. *British Journal of Psychology*, 101(1):47–68.

[Rochford, 1991] Rochford, L. (1991). Generating and screening new products ideas. *Industrial Marketing Management*, 20(4):287–296.

[Rogers, 2010] Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

[Romero, 2012] Romero, J. J. (2012). India's big bet on identity. *Spectrum, IEEE*, 49(3):48–56.

[Sabherwal and Chan, 2001] Sabherwal, R. and Chan, Y. E. (2001). Alignment between business and is strategies: a study of prospectors, analyzers, and defenders. *Information systems research*, 12(1):11–33.

[Sallam et al., 2011] Sallam, R. L., Richardson, J., Hagerty, J., and Hostmann, B. (2011). Magic quadrant for business intelligence platforms. *Gartner Group, Stamford, CT*.

[Samuelson and Zeckhauser, 1988] Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of risk and uncertainty*, 1(1):7–59.

[Schmidt and Calantone, 1998] Schmidt, J. B. and Calantone, R. J. (1998). Are really new product development projects harder to shut down? *Journal of Product Innovation Management*, 15(2):111–123.

[Schoonhoven et al., 1990] Schoonhoven, C. B., Eisenhardt, K. M., and Lyman, K. (1990). Speeding products to market: Waiting time to first product introduction in new firms. *Administrative Science Quarterly*, pages 177–207.

[Scott, 1987] Scott, W. (1987). Richard1987 organizations: Rational, natural, and open systems. *Englewood Cliffs, NJ: Pren-tice-Hall. ScottOrganizations: Rational, Natural, and Open Systems1987*.

[Shields, 1983] Shields, M. D. (1983). Effects of information supply and demand on judgment accuracy: evidence from corporate managers. *Accounting Review*, pages 284–303.

[Simon, 1972] Simon, H. A. (1972). Theories of bounded rationality. *Decision and organization*, 1:161–176.

[Simon, 1978] Simon, H. A. (1978). Rationality as process and as product of thought. *The American economic review*, pages 1–16.

[Smith and Alexander, 1988] Smith, D. K. and Alexander, R. C. (1988). *Fumbling the future: how Xerox invented, then ignored, the first personal computer*. W. Morrow, New York, 1st ed edition.

[Smith and Reinertsen, 1991] Smith, P. G. and Reinertsen, D. G. (1991). *Developing products in half the time*. Van Nostrand Reinhold New York.

[Soll, 2006] Soll, J. H. (2006). *Ideengenerierung mit Konsumenten im Internet*, volume 55. Springer DE.

[Srivastava, 2008] Srivastava, R. (2008). Changing retail scene in india. *International Journal of Retail & Distribution Management*, 36(9):714–721.

[Staw, 2009] Staw, B. M. (2009). Is group creativity reallyan oxymoron? some thoughts on bridging the cohesion–creativity divide. *Research on Managing Groups and Teams*, 12:311–323.

[Sternberg et al., 2003] Sternberg, R. J., Kaufman, J. C., and Pretz, J. E. (2003). A propulsion model of creative leadership. *The Leadership Quarterly*, 14(4):455–473.

[Sternberg and Lubart, 1996] Sternberg, R. J. and Lubart, T. I. (1996). Investing in creativity. *American psychologist*, 51(7):677.

[Subramanyam and Krishnan, 2003] Subramanyam, R. and Krishnan, M. S. (2003). Empirical analysis of ck metrics for object-oriented design complexity: Implications for software defects. *Software Engineering, IEEE Transactions on*, 29(4):297–310.

[Szajna, 1996] Szajna, B. (1996). Empirical evaluation of the revised technology acceptance model. *Management science*, 42(1):85–92.

[Tambe, 2013] Tambe, P. (2013). Big data investment, skills, and firm value.

[Teece et al., 1997] Teece, D. J., Pisano, G., and Shuen, A. (1997). Dynamic capabilities and strategic management.

[Te"eni, 2001] Te"eni, D. (2001). Review: A cognitive-affective model of organizational communication for designing it. *MIS quarterly*, 25(2):251–312.

[Timmons et al., 1987] Timmons, J. A., Muzyka, D. F., Stevenson, H. H., and Bygrave, W. D. (1987). Opportunity recognition: the core of entrepreneurship. *Frontiers of entrepreneurship research*, pages 109–123.

[Tuomi, 1999] Tuomi, I. (1999). Data is more than knowledge: Implications of the reversed knowledge hierarchy for knowledge management and organizational memory. In *Systems Sciences, 1999. HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on*, pages 12–pp. IEEE.

[Tversky, 1972] Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, 79(4):281.

[Tversky and Kahneman, 1973] Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2):207–232.

[Tversky and Kahneman, 1991] Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4):1039–1061.

[Tyebjee and Bruno, 1984] Tyebjee, T. T. and Bruno, A. V. (1984). A model of venture capitalist investment activity. *Management science*, 30(9):1051–1066.

[Urban,] Urban, G. L. *Design and marketing of new products*, volume 2.

[Utterback, 1994] Utterback, J. (1994). Mastering the dynamics of innovation: How companies can seize opportunities in the face of technological changeharvard business school press. *Boston, MA*.

[van der Aalst and Kumar, 2003] van der Aalst, W. M. and Kumar, A. (2003). Xml–based schema definition for support of interorganizational workflow. *Information Systems Research*, 14(1):23–46.

[Venkatesh and Davis, 2000] Venkatesh, V. and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management science*, 46(2):186–204.

[von Alan et al., 2004] von Alan, R. H., March, S. T., Park, J., and Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1):75–105.

[Wade and Hulland, 2004] Wade, M. and Hulland, J. (2004). Review: the resource-based view and information systems research: review, extension, and suggestions for future research. *MIS quarterly*, 28(1):107–142.

[Walls et al., 1992] Walls, J. G., Widmeyer, G. R., and El Sawy, O. A. (1992). Building an information system design theory for vigilant eis. *Information systems research*, 3(1):36–59.

[Walsh, 1995] Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization science*, 6(3):280–321.

[Watson and Wixom, 2007] Watson, H. J. and Wixom, B. H. (2007). The current state of business intelligence. *Computer*, 40(9):96–99.

[Weick, 1995a] Weick, K. E. (1995a). *Sensemaking in organizations*, volume 3. Sage.

[Weick, 1995b] Weick, K. E. (1995b). What theory is not, theorizing is. *Administrative Science Quarterly*, pages 385–390.

[Whitaker et al., 2010] Whitaker, J., Mithas, S., and Krishnan, M. S. (2010). Organizational learning and capabilities for onshore and offshore business process outsourcing. *Journal of Management Information Systems*, 27(3):11–42.

[Zheng et al., 2012] Zheng, Z., Fader, P., and Padmanabhan, B. (2012). From business intelligence to competitive intelligence: Inferring competitive measures using augmented site-centric data. *Information Systems Research*, 23(3-part-1):698–720.

[Zuboff, 1988] Zuboff, S. (1988). *In the age of the smart machine: The future of work and power*. Basic Books.

[Zucker, 1977] Zucker, L. G. (1977). The role of institutionalization in cultural persistence. *American sociological review*, pages 726–743.