PARDEE RAND GRADUATE SCHOOL

THE ARTS
CHILD POLICY
CIVIL JUSTICE
EDUCATION
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INTERNATIONAL AFFAIRS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
SUBSTANCE ABUSE
TERRORISM AND HOMELAND SECURITY
TRANSPORTATION AND INFRASTRUCTURE
WORKFORCE AND WORKPLACE

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

Jump down to document

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world.

Support RAND

Browse Books & Publications

Make a charitable contribution

For More Information

Visit RAND at www.rand.org

Explore Pardee RAND Graduate School

View document details

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAR 2006		**00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Multi-Attribute Strategy and Performance Architectures in R&D. The Case of The Balanced Scorecard	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Rand Corporation,Pardee RAND Graduate School ,1776 Main Street,Santa Monica,CA,90407-2138**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**310**	

This product is part of the Pardee RAND Graduate School (PRGS) dissertation series. PRGS dissertations are produced by graduate fellows of the Pardee RAND Graduate School, the world's leading producer of Ph.D.'s in policy analysis. The dissertation has been supervised, reviewed, and approved by the graduate fellow's faculty committee.

DISSERTATION

Multi-Attribute Strategy and Performance Architectures in R&D

The Case of The Balanced Scorecard

Athar Osama

This document was submitted as a dissertation in March, 2006 in partial fulfillment of the requirements of the doctoral degree in public policy analysis at the Pardee RAND Graduate School. The faculty committee that supervised and approved the dissertation consisted of Steven W. Popper (Chair), Bruce J. Held, Richard J. Hillestad, and Parry M. Norling.

PARDEE RAND GRADUATE SCHOOL

This dissertation is dedicated to my father, Lt. Col. (Retd.) Ahmad Osama Siddiqi (1947-2001,) whose faith in my abilities had no limits.

He did not live to see me through to the end of the tunnel,

but he never doubted the fact that I would,

one day, be there.

AND

My loving mother, Nusrat Osama Siddiqi, who is not only an inspirational parent and a wonderful teacher, but also a great friend.

The Pardee RAND Graduate School dissertation series reproduces dissertations that have been approved by the student's dissertation committee.

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from RAND.

Published 2006 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

PREFACE

This dissertation is about the alignment of strategic planning, performance measurement, and incentives systems within research and development (R&D) organizations. Specifically, it is an investigation into the appropriate use of one class of multi-attribute strategy and performance architectures, namely, the Balanced Scorecard, that has become a popular performance measurement and management framework. However, these have not been as well-received within research and development settings. This study takes a step back and asks the question: Are R&D organizations different? Do the underlying assumptions that make the Balanced Scorecard work hold true for research and development organizations? Do R&D organizations that the Balanced Scorecard realize performance "breakthroughs" that are often associated with framework? How might one modify or adapt the Balanced Scorecard framework before applying it to an R&D setting?

This study is submitted as a dissertation to the Frederick S. Pardee - RAND Graduate School for Policy Studies in March 2006 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Policy Analysis. The research was funded by RAND's Arroyo Center's Technology and Acquisitions Program from FY2002-2004 and by a generous gift from Don Conlan—a Board member at the Frederick S. Pardee – RAND Graduate School for Public Studies.

MULTI-ATTRIBUTE STRATEGY AND PERFORMANCE ARCHITECTURES IN R&D: THE CASE OF THE BALANCED SCORECARD

TABLE OF CONTENTS

i. Preface……..2

ii. Executive Summary……..5

iii. Acknowledgements……9

iv. Glossary of Abbreviations and Symbols etc……11

v. List of Figures……13

vi. List of Tables……....15

1. Performance Measurement in R&D: Background & Problem Definition……17

2. Introduction to Performance Measurement in R&D: A Review of Relevant Literature……29

3. The R&D Balanced Scorecard: Establishing a Reference Model ……..……58

4. The Research Questions and Methodology……106

5. Developing a Sense of the Universe: Lessons from a Survey of R&D Performers ……132

6. Strategy-Performance-Incentives (SPI) Systems in Practice: Lessons from Individual Case Studies……..164

7. Understanding Strategy-Performance-Incentives (SPI) Systems:
A Cross Case Comparison……205

8. Conclusions & Recommendations…….249

Appendices:

A. The Survey Questionnaire & Cover Letter……..262

B. Informed Consent Form……..271

C. Some Relevant Statistical Tables……274

Bibliographical References……..278

MULTI-ATTRIBUTE STRATEGY AND PERFORMANCE ARCHITECTURES IN R&D: THE CASE OF THE BALANCED SCORECARDS

EXECUTIVE SUMMARY:

Performance multi-dimensionality is an age-old problem. The notion that measurement of an organization's (or a system's) performance must incorporate, to the extent possible, all key dimensions has been discussed in a number of literatures, including, education, healthcare (McGlynn and Brook, 2001,) governance (Kauffman et al., 1999,) management, and measurement theory itself. In the business, management, and organizational contexts as well, performance multi-dimensionality is pervasive. The problem of performance multi-dimensionality—and hence measurement complexity—is most severe in research and development (R&D) settings due to the inherent multi-dimensionality of R&D's output and the long-term and intangible nature of the process itself.

One of the performance measurement approaches that internalizes the inherent multi-dimensionality of the organizational performance measurement challenge is the Balanced Scorecard (Kaplan and Norton, 1992, 1993, 1996, 2001, 2004 etc.) The key insight that triggered the idea of the Balanced Scorecard was the notion that organizational performance cannot be adequately measured by a single metric (or a single category of metrics) such as profit or financial metrics but must incorporate a whole series of metrics across a number of performance dimensions including input, process and output metrics, leading and lagging metrics, and metrics measuring tangible and intangible aspects of performance (Kaplan and Norton, 1992.) While the use of the Balanced Scorecard has spread in the private and non-profit sectors (Nevin, 2003), it remains, with the exception of Tippling et al. (1995) under-utilized and appreciated in R&D (Kerssens van-Drongelen, 1999.)

The attempts at doing so thus far have been preliminary, at best (e.g. Kerssens van-Drongelen et al., 2000; Bremser and Barsky, 2004; Loch and Tapper, 2002; Jordan et al., 2003; Jordan and Malone, 2003, etc.). This dissertation advances the state-of-the-art by asking the critical questions: Do R&D organizations satisfy the basic assumptions underlying the Balanced Scorecard? And does the adoption of Balanced Scorecard in R&D settings realize the kind of "breakthrough" improvements in organizational performance that are hypothesized by the founders of the Balanced Scorecard movement? The answers hold the key to unlocking the potential benefit of this framework in R&D settings.
We adopt a multi-pronged analytic strategy that builds upon the relative strengths of quantitative and qualitative methods. We conduct a mail survey of major US R&D

organizations across the entire organizational spectrum (i.e. public-sector, private-sector, and academic labs) to address the issues of prevalence. We conducted two mailings of a 6-part, 29-question survey instrument. While our response rates are low for this particular population, the results provide additional generalizability to the qualitative findings and help us in better interpreting the latter. We conducted an analysis of non-response bias by comparing the observable characteristics of early and late respondents (a la Omta, 1995) and find that our findings are more likely to be biased towards university labs and small public and private sector labs. We supplement this with a series of practitioner interviews (over 30 in all) with R&D directors, chief scientists, and lab managers to develop a broader understanding of the notions of strategy, performance, and incentives in R&D settings. Finally, we conduct six detailed case studies that attempt to understand organizations' strategy-making, performance measurement, and incentives systems within their unique organizational contexts. The case study analysis also aims to understand the quality of linkage between these activities and the improvements in performance that may be attributable to these activities or initiatives.

We took measures to improve the quality and generalizability of our qualitative findings (Yin, 1989, 1993.) For example, we used multiple methodological approaches and prior theory building to improve generalizability, adopted an experimental design approach to select our case study participants to ensure parsimony of effort, and developed interview and case study protocols to enhance standardization, reliability, and repeatability in data collection processes.

We test three sets of hypotheses dealing with the prevalence of multi-attribute performance measurement, in general, and Balanced Scorecards, in particular in R&D settings, especially focusing on various structural and derived features of the Balanced Scorecard methodology; the degree of balance, accessibility, participation, and transparency in the strategy-making process, especially focusing on whether or not R&D organizations take a differentiated (a la Treacy and Wiersema, 1995 and Porter, 1996) view of organizational strategy as necessitated by the Balanced Scorecard; and the quality of linkage between strategy-making, performance measurement, and incentives systems; its effect on the satisfaction reported from the performance measurement system; and its impact on organizational performance itself.

We find that the notion of performance multi-dimensionality is deeply ingrained in performance architectures of our respondents. Of the five pre-specified performance dimensions, our respondents, on average, found at-least 3.29 dimensions to be critical to

their overall performance. We also found the Balanced Scorecard to be much prevalent than originally perceived, accounting for 20% of all surveyed organizations followed by other frameworks like management by objectives (41%) and Total Quality Management (23%.) Cross-sectoral variations between public, private, and academic labs, however, do exist and are partially consistent with our theoretical priors.

While the qualitative analysis substantiates some of the above findings, it further clarifies and extends these in interesting ways. For example, we find the performance multi-dimensionality was an almost universal feature of organizations' performance measurement architecture irrespective of whether it was explicitly stated or not. Similarly, while a substantial number of organizations claimed to be using the Balanced Scorecard, an overwhelming majority did so in the name and not the spirit. Visibly absent were structural elements of the Balanced Scorecard such as cause-and-effect performance models and double-loop strategic learning.

We also found that balance, accessibility, transparency, and participation are increasingly being seen as desirable features of the strategy-making process and organizations, with minor exceptions, require participation of a wide cross section of researchers and scientists. These systems have been more successful when organizations tie real and tangible rewards and benefits with this participation. We also found that R&D organizations take a more diffused rather than differentiated view of strategy. We found a statistically significant portion of high performing organizations identified multiple simultaneous strategic foci (e.g. innovative leadership, customer responsiveness, and operational excellence) than those describing themselves as average performers. That the notion of "cheaper, faster, better" is an organizational reality in R&D deserves further investigation.

Finally, we found that R&D organizations do a poor job of linking together the strategy, performance, and incentives systems. R&D organizations, on average, also performed poorly on the consistency with which a strategic theme ran through these systems. In our qualitative analysis as well, we found several instances of multiple and conflicting strategic planning frameworks in organizations that led to confusion about what the strategy was and how it was being measured. While the implementation of the individual elements of the Balanced Scorecard did provide some anecdotal improvements in organizational performance and greater perceived satisfaction with the measurement system, we failed to detect "breakthrough" performance improvements from the implementation of the Balanced Scorecard.

The findings are useful for technology, innovation, and R&D management communities in a number of ways. While the notion that R&D and innovation was different from other forms of economic activity has been around, this is the first systematic attempt, to our knowledge, that makes this case in the context of implementing the Balanced Scorecard. Not only do R&D organizations not take a differentiated view of strategy—hence necessitating a modification of the Balanced Scorecard to suit the unique requirements of R&D organizations—but also the notions of performance and incentives may also be different in R&D settings thus requiring an adjustment of expectations from performance measurement approaches like the Balanced Scorecard. In many other ways, R&D organizations were not that different from other organizations. For example, we found evidence of the fact that R&D performance can be influenced, at-least at the margins, through better alignment of individual motivations and organizational goals. We also found considerable support for several other ideas that form the conceptual core of the Balanced Scorecard, at least individually, if not as a part of a package of organizational interventions.

Finally, and more generally, the study warns against selective adoption of certain features or elements in implementing an organizational management framework without thinking about its impact on the completeness of the overall framework. The study raises interesting possibilities for follow-on work, especially, leading to further clarifying the notions of strategy, performance, and incentives in R&D.

ACKNOWLEDGEMENTS

In the name of ***Allah*** (God)—the most Beneficent, the most Merciful—without whose will and mercy I wouldn't be writing this today. A large number of people—too many to be named individually—contributed to my lifelong learning that finds expression in this work. I would, however, identify two individuals who had the most defining influence on my own education and intellectual development. First and foremost, I owe my greatest debt to my mother, Nusrat Osama Siddiqi—a writer/journalist and an educator herself—who taught me how to write, both literally and metaphorically. Second, of the tens of teachers who played a formative part in my life, Nusrat Iqbal was the very special one. Among other things, she taught me how to read, and hence laid the foundations of the researcher in me. To these two special individuals and teachers, and many more from whom I have learnt throughout my life, I owe a thank you.

I also owe a special one to my wife, Ghazal, who, in addition to being a steadfast companion, was a source of unwavering support and optimism under the toughest of circumstances. Not only did she cheerfully tolerate my 16-hour workdays but also listened to my innumerable monologues about the Balanced Scorecard on dinner tables, car trips, and even amusement park visits. She and her parents were supportive and patient beyond the call of duty as I labored to bring this project to a close. They deserve a special acknowledgement for their optimism and belief.

Throughout my six years at the Frederick S. Pardee – RAND Graduate School and its parent institution, the RAND Corporation, I have met and had stimulating intellectual exchanges with a large number of very interesting and helpful people. Many of these were really thought leaders and icons in their own right and I immensely learnt from them. Of this distinguished body of gentlemen and women, my greatest gratitude goes to my Dissertation Committee members, Drs. Steven W. Popper, Richard J. Hillestad, Parry M. Norling, and Bruce J. Held. Had it not been for their faith in me and my "little" idea, this project would have never seen the light of the day. I am especially thankful to Steven Popper for giving me the confidence to take on this challenging topic at the very early stages of the process and believing in me all along. He went out of his way in providing whatever support and advice I requested. My thanks is also due to the excellent and committed faculty at the Frederick S. Pardee – RAND Graduate School who taught me the tools to undertake this task.

Several other individuals' contributions were key to making this work possible. First and foremost is the former Dean of the Pardee RAND - Graduate School, Dr. Robert Klitgaard, whose inspiring leadership and thoughtful mentoring was formative, to say the least. Bob was always there as a friend, in both letter and spirit, whenever I needed his help in the most difficult of times. Dr. Kenneth Horn, the former Program Director of the Arroyo Center's Technology and Acquisitions Program, and Bruce J. Held J.D., the current Director, and for most of my six years at RAND—my Principal Investigator and mentor—who not only got me started on this line of research but sustained it for over two years. Finally, a gift from Don Conlan—a Board Member of the Pardee – RAND Graduate School—came at a time when it was most needed. Without each of these individuals, I could not have done this work the way I was ultimately able to.

This list would be incomplete if I do not acknowledge the contributions of tens of individuals—scholars, R&D leaders, R&D managers, scientists, and engineers—who directly contributed to this work through their input as interviewees. I studied six organizations in great detail, interviewed practitioners and leaders from 32 other organizations, and had informal exchanges at conferences and hallways with many more. This work is my attempt to tell the story of their everyday reformatory struggles and creative endeavors. This is truly the fruit of our collective wisdom. I thank them for their confidence in me and this work, and for sharing their valuable thoughts and time.

An acknowledgement is also due to the wonderful support staff at RAND and the Pardee – RAND Graduate School. PRGS administration – Alex Duke, Rachel Swanger, Maggie Clay, and Kristen Copeland—were a source of support all along. Marcy Agmon, Margie Milrad, and Edith Nichols are wonderful ladies who knew how to make a new student feel at home at RAND and in America. The library staff at RAND, especially Gloria Dickey, was a tremendous help throughout my stay at RAND. Tora Bikson, Sandy Berry and Elliot Axelband provided very helpful comments on the survey and case study instruments. Several fellow students and RAND Colleagues provided comments during seminars. Craig Barela and Madeline Taylor helped mail out a thousand survey questionnaires and hence were critical, literally, to this project.

This work is the culmination of an educational experience that spans over 25 years. It is also the beginning of a process of life-long learning. Jacob Klerman—one of my favorite professors— once told us, "Research is done one brick at a time." I hope I have placed my own little brick in the wall of human knowledge.

GLOSSARY OF ABBREVIATIONS AND SYMBOLS

ABPA	Activity Based Profitability Analysis
AMC	Army Materiel Command
BSC	Balanced Scorecard
CS	Customer Satisfaction
COSEPUP	Committee on Science, Engineering, and Public Policy
CSA	Chief of Staff of the Army
CSF	Critical Success Factors
CSOx	Case Study Organization # x
DF	Derived Feature
DFSS	Design for Six Sigma
EMC	Employee Morale and Creativity
FPC	Financial Performance and Control
GPRA	Government Performance and Results Act
GDF	Generic Derived Feature
GSF	Generic Structural Feature
HBR	Harvard Business Review
IM	Innovation Management
IRI	Industrial Research Institute
KPI	Key Performance Indicators
LKM	Learning and Knowledge Management
MAPMS	Multi-attribute Performance Measurement System
MBO	Management by Objectives
MSC	Major Subordinate Command
NAP	National Academies Press
NAS	National Academy of Sciences

NGT	Nominal Group Technique
NRC	National Research Council
NRDEC	Natick Research, Development, and Engineering Center
OGSM	Objectives-Goals-Strategies-Measures
OMB	Office of Management of the Budget
PART	Program Assessment and Ratings Tool
PMS	Performance Measurement System
RDEC	Research Development and Engineering Center
RDECOM	Research, Development, and Engineering Command
RDO	Research and Development Organization
R&D	Research and Development
SF	Structural Feature
SOL	Standards of Leadership
SQF	Strategic Quality Framework
SRS	Strategic Readiness System
STEP	Science, Technology and Economic Policy
S&T	Science and Technology
TQM	Total Quality Management
TVP	Technology Value Pyramid
USACE	US Army Corps of Engineers
USAMRMC	US Army Medical Research and Materiel Command

LIST OF FIGURES

Fig #	Description	Page #
Figure 3.1	Performance Perspectives of a Balanced Scorecard	46
Figure 3.2	The Evolution of the Balanced Scorecard	52
Figure 3.3	Benchmarking the Balanced Scorecard	53
Figure 3.4	A Generic Balanced Scorecard Strategy Map	59
Figure 3.5	Graphical Representation of a Balanced Scorecard	63
Figure 3.6	A Modified Balanced Scorecard Template for Non-Profits Organizations	65
Figure 3.7	A Modified Balanced Scorecard Template for Public Sector Agencies	66
Figure 3.8	A Generic Performance Measurement and Management Framework	68
Figure 3.9	TQM's Philosophy, Assumptions, Principles, and Interventions	74
Figure 3.10	Example of a Balanced Scorecard for an R&D Organization	78
Figure 3.11	Dimensions of R&D Performance – Two Alternate Architectures	83
Figure 3.12	A Generic Strategy Map for an R&D Balanced Scorecard	85
Figure 4.1	The Multi-Pronged Research Methodology	101
Figure 4.2	Examples of Survey Research in the Population of Interest	103
Figure 4.3	R&D Performance and Incentives Survey Instrument	106
Figure 4.4a	Study Questions – Individual vs. Cross Case Analysis	109
Figure 4.4b	Data Collection Requirements and Interview Schedule	109
Figure 4.5	The Proposed Case Study Design and Analytic Plan	111
Figure 4.6	Theoretical Model / Priors for Case Study Analysis	112
Figure 4.7	Key Methodological Concerns Regarding Case Study Analysis	113

LIST OF FIGURES (CONTINUED)

Figure 5.1	Survey Item Assessing Emphasis on Performance Dimensions	131
Figure 6.1	OGSMs – Hierarchical Strategic Planning at CSO1	150
Figure 6.2	Balance in Strategy-Performance Systems like CSO1	153
Figure 6.3	Strategy-Performance-Incentives Alignment at CSO1	154
Figure 6.4	Performance Dashboards at CSO2	157
Figure 6.5	Balance in Strategy-Performance Systems at CSO2	159
Figure 6.6	Strategy-Performance-Incentives Alignment at CSO2	160
Figure 6.7	Performance Expectations at CSO3	164
Figure 6.8	Strategy-Performance-Incentives Alignment at CSO3	166
Figure 6.9	CSF and SQFs – The Strategic Planning Architecture at CSO4	169
Figure 6.10	CSO4's Performance Scorecard	170
Figure 6.11	Balance in CSO4's Strategy-Performance Architecture	172
Figure 6.12	Strategy-Performance-Incentives Alignment at CSO4	173
Figure 6.13	CSO5's Distinct Organizational Strategy	176
Figure 6.14	CSO5's Balanced Scorecard	177
Figure 6.15	Balance in CSO5's Strategy-Performance Architecture	179
Figure 6.16	Strategy-Performance-Incentives Alignment at CSO5	180
Figure 6.17	CSO6's Strategy-Performance Architecture	186
Figure 6.18	Strategy-Performance-Incentives Alignment at CSO6	187
Figure 7.1	Proposed Study Design and Analytic Plan	190
Figure 7.2	Theoretical Model and Priors for Case Study Analysis	196
Figure 7.3	Performance Multi-dimensionality – Cross Case Comparison	219
Figure 7.4	Alignment Between Strategy-Performance-Incentives Systems – A Cross Case Analysis	220

LIST OF TABLES

Table#	Description	Page #
Table 5.1	Statistical Snapshot of Responding Organizations	116
Table 5.2	Within-Group Response Rates – By Organization Type	117
Table 5.3	Observable Characteristics of Early and Late Responders	118
Table 5.4	Observable Characteristics of Various Respondent Types	119
Table 5.5	A Cross Section of Survey Respondents	121
Table 5.6	Selected Aggregate Organizational Structure Variables	122
Table 5.7	Breakdown of R&D Expenditure and Qualification Mix	123
Table 5.8	Key Performance Indicators of Surveyed R&D Organizations	125
Table 5.9	Types of Distinct Technology – R&D Strategies	127
Table 5.10	Are High-Performing R&D Organizations Strategically More Focused?	128
Table 5.11	Performance Dimensions Deemed Critical by R&D Organizations	129
Table 5.12	High Diligently Do RDOs Emphasize on Performance Multi-Dimensionality In their Strategic Management Systems?	130
Table 5.13	Potential and Actual Uses of Performance Measurement Systems and Information in R&D Organizations	132
Table 5.14	Types of Performance Measurement Systems or Frameworks Used	133
Table 5.15	Most Used (Top-Five) Metrics and Measurement Approaches for Individuals And Overall R&D Facilities	134
Table 5.16	Types of Compensation Rules Available a the Facility	135

LIST OF TABLES (CONTINUED)

Table 5.17	Types of Individual-level Incentives and Rewards Options Available	136
Table 5.18	Types of Laboratory-Level Incentives and Rewards Options Available	137
Table 5.19	Differences Between University, Public, and Private Sector R&D Labs on Key Dimensions of Relevance	140
Table 5.20	Differences Between High-Performing and Average-Performing R&D Labs On Key Dimensions of Relevance	144
Table 7.1	A Brief Snapshot of Case Study Organizations	192
Table 7.2	Salient Organizational Features of Case Study Organizations	193
Table 7.3	A Partial List of Organizations Interviewed	194
Table 7.4	Strategy-Making Among Case Study Organizations	200
Table 7.5	Performance Measurement Systems (PMS) Among Case Study Organizations	207
Table 7.6	Incentives Systems Among Case Study Organizations	214
Table 7.7	The Generic Performance Measurement and Management Framework and Case Study Organizations	222
Table C-1	Differences Between Research Labs and Development Labs	256
Table C-2	Differences Between Small and Large Labs	257
Table C-3	Differences Between Company R&D, Corporate and Dept. Labs	258
Table C-4	Differences Between World Class, Top Quartile, Above Average And Average Labs	259

CHAPTER # 1

PERFORMANCE MEASUREMENT IN R&D—BACKGROUND AND PROBLEM DEFINITION

Measuring the effectiveness of research and development (R&D) and innovation processes has long been a vexing problem faced by research and corporate managers alike. According to an Industrial Research Institute's (IRI) survey of its members, 'measuring and improving R&D productivity and effectiveness' was cited as highest among a ranking of 16 R&D related problems perceived by the members for two years, 1993 & 1995 (Ellis, 1997.) Traditionally R&D was seen as a rather 'chancy' process, dependent upon a lot of idiosyncratic factors that went unmeasured partly due to a popular recognition of the uncertainties involved and the highly fuzzy and ill-defined nature of its output (Baglieri et al., 2001.) It was commonly believed and widely accepted, therefore, that little could be done to improve the efficiency of R&D.

Another factor that helped R&D escape close scrutiny through much of the known history was the charisma associated with the work of inventors and scientists. The popular veneration of technology and its progress ensured that public opinion towards technology was based more on faith than on systematic appraisals of its benefits and shortcomings (Geisler, 2000: p. 13). Consequently, assessing the productivity of an R&D organization (RDO) was considered more a matter of qualitative judgment depending on assessor's ability to 'sense' whether a research organization was more productive than another (Fusfeld & Langlois, 1982). With changing times and the perception of the R&D process, the focus has shifted to trying to improve the effectiveness of research & development.

The measurement problem faced by the R&D community today is quite complex. As R&D becomes a more and more complex (Rycroft et al., 1999) and a multi-disciplinary activity, so does the R&D organization. Today's R&D organizations are pressed by the demands of efficiency, effectiveness and accountability on the one hand, and the requirements for greater multi-disciplinary and multi-site collaboration on the other; all of which takes place in an environment of increased technological discontinuities, accelerated technological obsolescence (Moore, 1965) and competence erosion, and enhanced competition. This makes the problem of creating such organizations and measuring their performance, an especially difficult one.

Concurrent to the above development, increasingly, R&D leaders and managers are coming to terms with the fact that the days of almost unrestricted ("no questions asked")

sponsorship of research—that started somewhere around the WWII—are over (Kostoff, 2000). Instead, today's R&D organizations, both in the public and corporate sectors, are being held to a much higher standard of accountability than they ever were and are required to justify their value to the larger organization (at times, even the society) of which they are a part. Measurement of R&D—whether for the purpose of assessing the health and continually improving the R&D enterprise or justifying its value to major stakeholders—is central to this discourse.

1.1—THE PRIMACY OF PERFORMANCE MEASUREMENT IN R&D

At the broader national innovation system level, the fundamental questions that are at the heart of the philosophical debate on research sponsorship are increasingly coming under intense debate between research managers, policy makers, bureaucrats and the technical community. These include, among others: what sort of research should the government sponsor? Is there a case for private sector sponsorship of basic research? Should the government get involved in commercialization of research? Should the government maintain research capacity in fast obsolescing areas? At the level of an individual R&D organization as well, the more pragmatic issues are subjects of an ongoing discourse within the R&D management community. These include, among others: what constitutes "success" in the context of an R&D organization? How to manage R&D in organizations better so as to make R&D, what was once thought of as a chancy process, more predictable and planned? How to better do incremental and discontinuous innovation, and do both at the same time? How to bring better tools and techniques to bear on areas like R&D project selection and technical risk assessment? How to incentivise performance and reward success?.

Both of the above debates—the philosophical, and the pragmatic—have lives of their own and play out almost simultaneously, albeit with different timeframes and tempos, and within overlapping yet different audiences. They are, however, far from being settled. In fact, one can argue that they are ongoing and never ending. Their intensity, importance, and relevance, however, are affected in times of paradigm shifts within and external to the R&D enterprise. Quite similar to the Kuhnian scientific revolutions[1] that affect the substance of scientific research in various fields of scientific discovery, the conduct of scientific research itself—a field that is sometimes termed as "research on research"[2] (IRI,

[1] Documented in Thomas Kuhn's classic "Structure of Scientific Revolutions" (Kuhn, 1962),

[2] The Industrial Research Institute (IRI), for example, has a Research-on-Research (ROR) committee since 1971 as IRI's sponsoring and leadership body for improvement of R&D leadership and

1971)—is also subject to similar cycles of enlightenment and agreement followed by disillusionment, chaos and then all over again.

For example, the philosophical debate on research sponsorship and management was going through a period of a "comfortable" agreement—the "normal science" in Kuhn's terminology—during much of the cold war where the benefits of research, both basic and applied, were adequately expressed and abundantly clear to the various stakeholders, the case for public support of R&D had been explicitly made and agreed upon (a la Vannevar Bush's Endless Frontier[3]) and industry—enjoying an era of unprecedented world leadership—was pouring billions of dollar in R&D of both the basic and applied kinds. The fundamentals that were driving this consensus or agreement, however, changed with the end of the cold war and soon the R&D community found itself in the middle of disillusionment and chaos—a transition to crisis as Kuhn would put it.

For over a decade and a half now, the R&D community has faced an in-equilibrium of sorts, whereby it is increasingly called upon to justify its relevance and value to the society. This identity crisis, or "blurring of the paradigm" in Kuhn's parlance, is felt more in some areas and sectors than others. An apt example would be the confusion in the post-cold war era over the role of the public R&D in the larger society. With the threat of the "evil" communist empire gone—and along with it the national security-based raison d'etre—the consensus on what public-sector funded R&D must accomplish has shifted. The 1990s saw a greater thrust toward commercialization of publicly-funded R&D—something that was quite out-of-favor in the earlier consensus. Needless to say, however, that like Khunian scientific revolutions, the cycle repeats itself—and indication of which is the post-9/11 re-evaluation and re-alignment in the America's priorities and the new found role of public sector R&D.

On the more pragmatic front as well, the R&D enterprise goes through similar periods of enlightenment, agreement, disillusionment and crisis whereby the industry is galvanized by a new management paradigm or analytic tool that promises to provide a fix to all previously unanswered questions and problems. This is followed by enthusiasm for this

---

management practices. ROR has been an active body within IRI and the latter's website describes several communities and a portfolio of completed and current projects funded and executed by the committee. For details visit:
http://www.iriinc.org/Content/NavigationMenu/Collaboration_and_Research/Research_on_Research/Research_on_Research.htm

[3] See Bush (1945)

paradigm (e.g. Re-engineering, TQM, Design for Six Sigma,) its wholesale adoption, a period of agreement followed by disillusionment and then abandonment in the favor of a new and different paradigm. In due course "normal" operations of each paradigm, however, many more and newer questions are raised than were settled earlier and new evidence emerges that point toward its ineffectiveness in solving the organizational and management problems that it was supposed to address. These management paradigms—or fads, as they are often called for the zeal with which they are initially followed and later abandoned—normally first emerge in the business community, then government, and finally education sectors (Birnbuam, 2000). The R&D enterprise is no exception.

The philosophical and the pragmatic debates are far from being over, and depending on how much one believes in their Kuhnian character, would perhaps go on forever. There is, however, one critical element that ties both these otherwise parallel yet related debates together, namely, the science and the art of R&D measurement. The ability to measure R&D—not only in broad philosophical terms like its private and social benefits (or spillovers) to the society but also in pragmatic terms such as the health of an R&D system (e.g. a lab, a series of labs, or an entire national innovation system) and its ability to achieve its stated objectives gets at the heart of the entire philosophical and pragmatic debates on R&D management, organization, and sponsorship.

Measurement—in the conduct of research, as it is in the substance of research itself—has a tool-like character. While the techniques of measurement might improve as well as more constructs to be measured added over time—as has been the case with substance of research itself—the fundamental idea of measurement remains stable and ties the various generations of research together. For example, adequately developed R&D measurement systems allow labs to monitor their performance over time, identify bottlenecks and avenues for improvement (the pragmatic debate), and justify their presence to the wider corporate community and the society (the philosophical debate). While the nature and meaning of "success" and "performance" might change over time, from one generation of R&D to the next, a well-developed and stable R&D measurement paradigm would continue to serve the R&D establishment over time. It is this centrality to the two debates and commonality across the changing paradigms over time that makes R&D measurement systems so critical to the health of the R&D establishments worldwide.

A well-developed and stable R&D measurement paradigm provides us with the ability to clarify things, communicate across paradigms, and facilitate the transition between them.

This concept of performance measurement—in its pragmatic, organizational context, but with ramifications for the larger philosopher debate—is the purpose of this inquiry.

1.2—THE BACKGROUND AND MOTIVATION FOR THE RESEARCH

This research was conceived in the backdrop of several related developments the most important of which was a paradigm shift in the sponsorship and conduct of public R&D in the United States. In the after-math of the 9/11 and the ensuing "War on Terrorism", the Department of the Army—on the initiative of Secretary of Defense Mr. Donald Rumsfeld—embarked upon a re-organization exercise as a part of the larger re-organization of the Pentagon and Department of Defense. As this movement for re-organization progressed and took root, the US Army Materiel Command which then housed a major portion of Army R&D establishment (i.e. Army Research Office, Army Research Lab, and several Research, Development, and Engineering Centers) engaged RAND Arroyo Center to help in this re-organization process.

One of the options under consideration was to spawn the R&D assets of the US Army into a separate US Army RD&E Command (RDECOM) from their structure under the Army Materiel Command's Major Subordinate Command (MSC) called Laboratory Command (LABCOM.) This, it was hypothesized, would free the R&D establishment from the conflict of interests under the current arrangement and provide it with the leadership and autonomy to pursue its mission in support of the war-fighter. RAND's task was to help the Army R&D establishment think through and clarify its objectives overall and from the re-organization process, suggest alternative models for the new Army RDECOM and make recommendations on how to best structure and carry out the re-organization process itself. RAND's research team regularly visited and attended meetings and contributed in the strategic planning and assessment phases of this exercise.

As a part of this engagement, RAND undertook a historical study of the US Army R&D establishment looking at the evolution of the Army R&D organization since the year 1800 till the present and identifying the key drivers of various re-organization efforts that were undertaken during the intervening 200 year period (Wong, 2003). In addition to providing the Army leadership with a long-range perspective on the on-going re-organization effort, several important findings came out of this historical analysis. The most important of these findings was that over the 200 year old history of the organization—during which it went through 5-6 major re-organizations—every time the Army R&D organization went through

with a process of re-organization, a similar set of concerns or themes were cited as reasons for doing so. These included: achieving better control of money, developing better working relationships between various labs, doing integrated research to facilitate transitions, and minimizing the duplication of research through better visibility.

What was even more striking was that every new re-organization effort was premised on the failure of the former to rid the organization of these issues without presenting actual data and performance information to support the claim. Thus while the impetus for re-organization came from the Army leadership, the people in ranks expressed ignorance of what was wrong with the current structure and how the new structure was going to improve it. Questions like: What was precisely wrong with the current organization, defined in a concrete and verifiable manner? What was the true ideal of what Army R&D organization needs to be? And how would the new organization help improve the status-quo? were answered more on the basis of gut-feeling rather than any concrete evidence and data. The re-organization debate seriously lacked a common language—the performance measurement philosophy—that could link the various organizational paradigms and across different time periods. A common language could transform the re-organization effort into one attempting to the same things all-over-again into one that was a concerted effort to improve the performance of Army R&D lab system.

In parallel to these, however, there were certain other developments as well. The most important of these developments was a performance measurement or strategic management framework called the Balanced Scorecard. The Balanced Scorecard occupies a central position in the President George W. Bush's Management Agenda (Whittaker, 2003.) Needless to say that the culture of performance scorecards and spotlight scoring systems that began then has permeated the management philosophy and reporting of the Federal Government[4]. One Washington insider, Steve Rohleder, the Managing Partner of Accenture's USA Government Practice in Washington, was quoted in Industry Week as expecting a major shift in the government philosophy towards performance and results from its previous "science-project" level to that of business operations. "If these guys really want to operate like a business, when they go to Congress to justify their appropriations, they will tie them to outcomes -- and they will explain . . . how they're going to measure those outcomes and report back progress against them" (McClenahen, 2001).

Following up the tone set by the Bush Administration's emphasis on the use of the Balanced Scorecard in the public sector agencies, the Army leadership also embarked upon implementing a Balanced Scorecard across the entire administrative and operational

[4] Please see http://www.whitehouse.gov/results/agenda/scorecard.html for an extensive example.

spectrum of the US Army. Gen. Shinseki directed his Deputy Chief for Operations and Plans "to develop a readiness reporting system that is accurate, objective and timely in its measurement of the Army's ability to support the National Military Strategy and allows the CSA to direct resources in order to influence readiness across the Army."[5] A Report from the Secretary of Defense describes the centrality of the Balanced Scorecard to US Army's reporting and measurement strategy in the following manner:

"[U.S. Army's] Strategic Readiness System (SRS) is based on the Balanced Scorecard (BSC) methodology, which was designed to provide Army leadership with a single system that communicates our mission, vision, strategic objectives and priorities. It enables the Army to monitor progress against our vision, to forecast strategic performance and to make adjustments, as necessary, to resources, personnel and policy…Individual commands, in addition to our leadership, also have extracted tremendous value from the BSC. Units have been able to meet their mission essential task lists more effectively and to focus simultaneously on readiness and overall transformation toward the Future Force. The SRS balanced scorecard allows each team to evaluate recent unit performance in a way that cuts across organizational "silos" (e.g. logistics, operations, medical, training and other staff areas). People from different organizations within the Army have easy access to scorecard data, and can thus quickly come together to work common issues." (Rumsfeld, 2004)

Over the years, the Balanced Scorecard has been developed and rolled down throughout the Army's organizational structure. According to a sub-contractor's estimate, over 400 SRS Balanced Scorecards are planned to be rolled out to 40,000 individuals and formations within the Army (CorVu, 2003). Several of these scorecards, including Army's overall scorecard, are in the public domain. US Army Medical Research and Materiel Command (USAMRMC) and US Army Corps of Engineers (USACE) are among these formations. The newly materializing Army Research Development and Engineering Command (RDECOM) was also tasked with putting in place a performance measurement system based on the Balanced Scorecard approach. The RAND research team supporting the exercise saw this as an opportunity for a contribution for two reasons. Firstly, as discussed above, RAND had already identified the lack of a common performance measurement architecture as one of the weaknesses that hindered the ability of the Army R&D lab to reform for the better (i.e. not having a common language across paradigm shifts, it did not know, in precise and quantifiable terms, what the term "better" meant). Secondly, the coincidence of a re-organization and the implementation of a new strategic planning and performance system

[5] CorVu Corporation, 2005 at:
http://www.corvu.com/documents/casestudies/CorVu%20US%20Army%20Success%20Story.pdf

allowed the possibility of by-passing the usual organizational resistance and inertia that often precedes deployment of Balanced Scorecards in well-established and mature organizations.

The RDECOM's interim leadership[6], however, was mildly skeptical of the Balanced Scorecard approach. While the Balanced Scorecard had considerable following and some anecdotal evidence of success in the corporate community, it was not particularly well-received within the R&D community. Few readily identifiable success stories of R&D organizations implementing a Balanced Scorecard existed and even fewer reported any dramatic improvements in performance or employee morale. In fact, such attempts were often seen with skepticism by the R&D community—the rank-and-file bench scientists and middle managers—and were fiercely resisted. One such example was the Industrial Research Institute's (IRI) attempt to develop and popularize its own version of the Balanced Scorecard (Tipping et al., 1995.,) The Technology Value Pyramid (TVP) had only received a modest degree of success within the R&D community. The question, then, for the RDECOM was whether to subscribe to the Army-wide Balanced Scorecard initiative and develop one of its own, or take guidance from the lack of response of the R&D community and opt for a different performance architecture.

Performance multi-dimensionality—a concept which is at the heart of the Balanced Scorecards approach—is not foreign to the Army R&D establishment. The US Army Natick Research and Development Center (NRDEC) was part of major National Research Council (NRC) study titled "World Class Research and Development" (NRC, 1996) that identified five pillars (or dimensions) of world-class research and development organizations, namely, resources and capabilities, strategic vision, quality focus, customer focus, and value creation. Natick RDEC's mission is to "develop products that maximize soldier survivability, sustainability, mobility, combat effectiveness, and quality of life". The study was commissioned to provide NRDEC's leadership with a framework to think about and assess itself vis-à-vis an ideal "world-class" R&D facility. Following the study, NRDEC implemented its recommendations developing an organization-wide strategic planning and performance measurement architecture that endorsed a multi-dimensional view of organizational performance. Following the implementation of the recommendations, NRC's Standing Committee on Program and Technical Review of US Army Natick RDEC ("the Natick Standing Committee") assessed the Natick RDEC's progress and achievements

[6] The RDECOM was still in the flux of transition from AMC's Major Subordinate Command (LABCOM) to a separate command itself and, at the time of RAND's engagement, was headed by a General Officer in an interim or transitional capacity.

against the characteristics and metrics identified in the earlier report. This assessment is documented in NRC(1997).

In its essence, NRDEC's adoption of the five-dimensional model of performance assessment was RDECOM's first foray into a Balanced Scorecard-type performance assessment architecture, minus the strategic planning framework of a Balanced Scorecard. NRDEC continues to be highly rated and assessed for its organization and performance within the U.S Army Lab system. In 2004, for instance, it was rated for three out of four consecutive years as the Department of the Army Research and Development of the Year Award in Small Development Lab Category. The citation mentioned NRDEC's contribution to its mission and vision of soldier support, mirroring the performance dimensions identified in President's Management Agenda while acknowledging the role of NRDEC's Balanced Scorecard in tying the two together. It states:

"This year, many of the criteria chosen for evaluation were directly linked to the president's management agenda. The soldier center used its balanced scorecard strategic map to link resources, human capital, business practices and strategic initiatives to its core competencies of warrior-related technology generation, application and transition; warrior systems technology integration and transition; and solving warrior and warrior support-related field problems rapidly." [7].

Another concern that we (RAND research team) had, as we attended the meetings of RDECOM leaders and planners that deliberated on the issues of re-organization, strategic planning, and performance architecture, was that of "thoughtless" adoption of the Balanced Scorecard (or any alternative) methodology on the newly formed organization. Not only would such an implementation fail to realize the benefits of streamlined and effective strategic management architecture but also result in the loss of an opportunity that had arisen due to the confluence of organizational re-design and implementation of a performance measurement architecture.

As the time for materializing the RDECOM approached and the tendency to engage in a political rather than an organizational process became stronger, the chances of the above concern materializing grew significantly. We began looking for a systematic approach to assessing the appropriateness of implementing a Balanced Scorecard-type performance architecture in RDECOM while at the same time pondering over the larger philosophical

[7] For detailed citation, please see: http://www.defenselink.mil/transformation/articles/2004-11/ta111804b.html

issues of designing effective performance measurement and strategic planning architectures in R&D.

1.3—THE SCOPE OF THE RESEARCH

The notion of performance measurement in R&D creates the imagery of several kinds in attentive and informed minds—e.g. cost-benefit and ROI analyses, bibliometric analyses of patents, publishing, and citation counts, scientific peer reviews, and sociological studies of scientists and scientific institutions. Each of these constructs represents a real and worthy goal for an analytic exercise. Each of these types of studies also definitely has a role and a contribution in our overall understanding of the scientific research and development enterprise and must be subjects of separate studies in and of themselves. When the mention of a phrase invites so many different connotations, meanings, and imageries, it is important to clarify the scope of the work being undertaken in an upfront and explicit manner.

While the proposed study would have several commonalities with other similar studies that fall under the rubric of R&D performance measurement, it would depart from these in significant ways. For example, the starting point for this research, like any other research in this area, is a deep understanding of the R&D process, in general, and performance measurement in R&D, in particular. This would include, not only an understanding and appreciation of the philosophical issues associated with notions of performance measurement in R&D but also an awareness and appreciation of the "performance analyst's toolkit" comprising various methodologies and tools, their appropriate usage, their strengths, biases, and weaknesses. These tools and methodologies form the various components and sub-components of an effective measurement system for R&D performance.

Due to the broad nature of this investigation, we would attempt to scope this research by use of the negatives i.e. what it would not attempt to do rather than what it would do. First, the proposed project would not seek to make a contribution in the area of assessing or calculating the economic benefit or value of R&D and R&D related spillovers etc. This is a fairly mature enterprise and the interested readers are advised to refer to the work of other luminaries (e.g. Mansfield, Griliches, Nelson). Second, the proposed project would not try to look at organizational redesign as a distinct performance improvement option in itself. It would, however, entertain re-design within the context of the implementation of a performance measurement system. Third, this study would not, with the exception of the

Balanced Scorecard itself, delve into the micro-level specificities of the various tools, techniques, and methodologies for measuring R&D performance and impacts. Fourth, this study would not delve into the issues of individual performance per se, except however, as an attempt to explain the linkage between individual and organizational performance.

Having clearly discussed what is not in the scope of this research, let's discuss what is within the scope of the proposed research. The research would have two major parts. The first part of the proposed project would comprise a theoretical and empirical inquiry into performance measurement in R&D.

On the theoretical side, we would build upon an understanding of established literature on performance measurement, in general, and R&D performance measurement, in particular, and develop a generic model of a Balanced Scorecard-type performance measurement system to guide our discussion. This task would require an understanding of various tools, techniques, methodologies, and candidate performance measurement systems that have been (or can be) used in various organizational settings, including but not limited to, Management-by-Objectives (MBO), Total Quality Management (TQM), Design for Six Sigma (DFSS), Technology Value Pyramid (TVP), and the Balanced Scorecard (BSC). It would focus on an organizational view of performance that answers the question: how can we assess the overall health of an R&D organization? To that end, the study would try to survey the Balanced Scorecard literature and develop a reference model that can be used to as a benchmark the empirical analysis.

On the empirical side, we would attempt to survey the actual practice of performance measurement within various types of R&D environments (e.g. academia, public sector labs, and corporate labs as well as basic and applied research and development.) Questions of the nature: How do various types of R&D organizations think about and measure performance? What are the most prevalent performance measurement approaches and philosophies and what are the underlying assumptions that make them work? How is the performance information being used in various types of R&D organizations? And how do organizations link their performance measurement architectures with the wider strategic management systems (e.g. strategic planning, and incentives systems). We would also attempt to empirically test the various assumptions behind the generic model of Balanced Scorecard in R&D.

The basic insight behind the Balanced Scorecard approach to performance measurement is as follows: Performance, it is sometimes asserted, is not a uni-dimensional construct and

trying to measure it through one type of metrics risks missing other important dimensions of performance. This calls for a holistic approach towards organizational development and assessment whereby the demands of short-term financial performance are balanced against those for sustaining long-run innovation.

Over the years, however, the Balanced Scorecard has evolved into a full-fledged strategic management system[8] with performance measurement as its centerpiece, and has been used fairly successfully, at least at an anecdotal level, within corporate, public, non-profit, and healthcare sectors. Its use, however, has been rather limited within the R&D enterprise. One attempt to promote a Balanced Scorecard-type approach within R&D is the Technology Value Pyramid (TVP) designed and championed by one of the largest and most prestigious associations of commercial R&D labs in the US—the Industrial Research Institute (IRI)—as an R&D equivalent of the Balanced Scorecard (Tipping et al., 1996)[9]. For largely unexplained, although widely speculated, reasons the Technology Value Pyramid failed to gain the kind of acceptance and adoption within the R&D community that its intellectual father, the Balanced Scorecard, was able to achieve within the corporate world. Empirical accounts of the implementation of technology value pyramid are virtually non-existent in the R&D practitioner literature.

Incidentally, though, some qualities of multi-attribute performance measurement systems lend themselves very well to the R&D setting. For example, the roots of the Balanced Scorecard itself lie in the need to measure and monitor intangibles and long-term performance drivers in business organizations. These tensions between the long and short term, as well as between the quantifiable and non-quantifiable, are most conspicuous in research and development (R&D) where the product has a dual character (i.e. knowledge and gadgets) and investments are often made "on faith" with the expectation of returns and outcomes years down the road. Despite these similarities and overlaps, however, the rather limited use of Balanced Scorecard has remained an intellectual curiosity—one that has been a primary motivation for this research.

---

[8] Several Publications by Dr. Robert Kaplan and Peter Norton discuss that possibility. Also, see Osama, Athar et al., "Designed to Deliver: Using a Systematic Organizational Framework to Create Army's New RDE Organization from First Principles", DRR-3000-A, RAND, March, 2003, for a discussion on the use of Balanced Scorecard as an OD Tool.

[9] This point was also alluded to during the author's telephone interview/conversation with Dr. Alan Fusfeld—one of key figures in the IRI's development of Technology Value Pyramid and a co-author of Tipping et al., 1995.

Consequently, the study would look at the appropriateness of using a multi-attribute performance assessment system—namely, the Balanced Scorecard—in the R&D environment. In doing so, it would explore questions like: Do R&D leaders and managers think about R&D performance as a multi-dimensional construct? How do they attempt to weigh (or achieve balance) across the multiple dimensions of R&D performance? Do the underlying assumptions of Balanced Scorecard-type performance measurement architectures hold true for organizations engaged in R&D? What are some of the mechanisms through which R&D organizations achieve improvements in performance and how do they relate with the Balanced Scorecard's modus operandi for the same? What benefits do R&D organizations using multi-attribute performance assessment systems realize as compared to non-R&D organizations using similar systems or R&D organizations using other kinds of systems?

Understanding these questions, in the light of theory and practice of R&D measurement, would provide guidance and insights into whether or not, and how to structure a multi-attribute performance measurement system for an R&D lab system. For example: How many (and what type of) metrics are needed to assess the performance of an R&D lab (or system of labs)? How closely do these metrics need to mirror the organizational strategy and how best to bring about this alignment between organizational strategy and performance measurement? How should one pick and validate metrics for measuring R&D? What are the best uses of such metrics, keeping in mind the possibilities of dysfunctional behavior, and what we learn from the experience of organizations that have achieved desired improvements in R&D performance? What constitutes "appropriate expectations" from setting up a performance measurement system in within R&D? What might be the best use of the information provided by such a system? What sort of implementation bottlenecks can one expect during the course of such an implementation and how might one prepare for these? How might the validity of such systems and the information provided by them be assessed?

While the proposed research is framed in the context of the US Army's RD&E Command—a public sector RDO—it would, we believe, have general applicability to a whole spectrum of R&D labs and organizations, both in public and private sectors. A carefully selected set of private and public sector case studies and an in-depth review of literature on performance measurement would allow us to make generalized conclusions that would far transcend the Army RDE Command. Also, while the proposed research will focus on performance measurement in the context of an R&D lab, the techniques explored and developed can be scaled up or down and translated to do the same for any R&D system

such as a research team, a virtual organization, a group of collaborating institutions, a system of labs, or the entire national innovation system.

In the next chapter, we would look at the evolution and acceptance of the concept of performance measurement in the research and development enterprise (chapter-2) in an attempt to lay the groundwork for a more in-depth assessment of the Balanced Scorecard methodology (chapter-3) and its subsequent examination with a view towards applying it to the research and development enterprise.

CHAPTER # 2

INTRODUCTION TO PERFORMANCE MEASUREMENT IN R&D—A REVIEW OF RELEVANT LITERATURE

The concept of performance—and hence measurement—in R&D has evolved considerably over the last half of the twentieth century reflecting the changing paradigms, both in organizational and sponsorships realms, that have dominated the conduct of research over the last five decades. In the past, for example, the popular support for the "seemingly mysterious activities of a small group of inventors" ensured that "there was little in the form of reliable measurement of technology's outcomes" (Geisler, 2000: p. 13). With the passage of time, however, as science and technology has become an increasingly organized activity, and its influence has increased, so has the need and demand for greater public scrutiny. In order to understand these shifting moods and their effects on the concept of performance in R&D, it would be a useful exercise to briefly review the history of research and development in the United States.

2.1—R&D MEASUREMENT IN HISTORICAL CONTEXT

Kostoff (1997)—in a common sense analysis of the evolution of the research and development enterprise through history—talks about the successive phases of evolution of basic research from a 'rich man's pastime' to industrial sponsorship to government support. During the inter-war period, research funds in the world were fairly limited and European research had an advantage over the rest of the world primarily due to a head start and other historical factors that prevailed in the pre-war era. The balance of power shifted after the World War-II with much of Europe and Asia undergoing massive rebuilding efforts due to the destruction of the Great War. This was the time of the rise of industrial sponsorship in the US. The US had virtually no competition from around the world and, therefore, American companies could divert substantial resources into research and development.

As the time passed and European and Asian corporations came out of their post war rebuilding as strong, if not equal, competitors to the US companies, profits margins shrank and so did the industrial R&D budgets. Many companies, operating in highly competitive environments and looking for short-term profitability, could no longer afford the luxury of investing heavy sums of money in basic research that formed the foundation of the innovation pyramid. Unable to make or even foresee enough profits from undertaking basic

research, the world saw huge cutbacks in industrial support of R&D during the seventies and the eighties—in real if not nominal terms.

With support from the corporate coffers dwindling, the proponents of big "R" & small "d" invoked the classical 'public good' argument to attract government support for their activities. This coincided perfectly with the fast popularizing notion of "Big Government" in the US and elsewhere. In the post war period, leading through to the end of the cold war, the federal government in the US became a major supporter of basic (and applied) research. A similar shift occurred in Europe and Asia with a time lag that accounted for the post-war reconstruction. One of the reasons advanced for this massive growth in public investment in basic and applied research was the national security argument aimed at winning the technology race from the former Soviet Union and to check the growth of the "evil" communist empire around the world.

However, by the end of 1980s and early 1990s, government priorities shifted, once again, under increasing pressure to maintain fiscal discipline and less intervention. As governments tried to become 'smaller', R&D had to face competition for funding support from a host of social and other causes. The fall of the Soviet Union further aided this trend. With the national security imperative all but gone, the stage was set for the current move towards greater accountability and efficiency in research and development. These developments also had an impact on how performance was perceived and measured in the R&D community or demanded from it by external stakeholders (e.g. corporate leadership, congress and other public sector entities.)

Expressing similar themes from a different standpoint, Geisler (2000) emphasizes the formalization and institutionalization of the R&D enterprise in the western societies as providing the most important impetus for measurement and evaluation. It puts forth an interesting analysis of the emergence of institutions in the S&T domain, such as, universities, public laboratories, industrial labs, and technology districts etc. "Once the trend of institution-building reached a comfortable level of growth", it contends, "science and technology is now faced not with the question: 'Is it good or necessary for us to have it?' but with a different set of questions…'How can we improve its performance?' and 'What exactly do we get from it?'"

We start our analysis of the evolution of the concept of performance in R&D right from the closing days of the second world war by looking at the event that set the stage for expansion of public support for science and R&D throughout the second half of twentieth

century, namely, the release of Vannevar Bush's "Science: The Endless Frontier" report in July 1945. While the report called for a significant shift in the involvement of the federal government in the country's post-war S&T establishment—citing the national objectives of security, health, economy and knowledge creation as motivations to do so--it was not totally oblivious of the ideas of performance (Bush, 1945). On a number of occasions, the committees charged with putting together the various sections of the report talked about effectiveness being an overriding concern. For example, in a number of instances the report considers constructs like costs of recommended policies, costs of inactions, effectiveness and efficiency, trade-offs between alternatives, and public welfare and private profits originating from public R&D investments etc.

Several sections of the report implicitly address the notions of performance and utility. In a discussion on means and ends of research, the report supports public investment in applied research in the following manner: "The scientist doing basic research may not at all be interested in the practical applications of his work, yet further progress of industrial development would eventually stagnate if basic scientific research were long neglected". At another instance, the report talks about the idea of proportionality between promoting multiple dimensions of knowledge in the following words: "It would be a folly to set up a program under which research in the natural sciences and medicine was expanded at the cost of the social sciences, humanities, and other studies so essential to national well-being".

As the members of these committees carefully deliberated upon issues of cost and benefit from public sector research, they also acknowledged that the cause was of such importance and the precise assessment and fine-tuning of the system so difficult that some errors may be accepted as necessary costs of doing the business of research. In the concluding chapter the report mirrors this sentiment: "We have given much thought to the question of how plans for the use of Federal funds may be arranged so that such funds will not drive out of the picture, funds from local governments, foundations, and private donors. We believe that our proposals will minimize that effect, but we do not think that it can be completely avoided. We submit, however, that the nation's need for more and better scientific research is such that the risk must be accepted."

We highlight these sections to illustrate the fact that ideas of performance and effectiveness—and hence measurement—were not foreign to at-least some sections of the well-meaning scientific community—several people serving on Vannevar Bush's committees among them—even as back as 1945 and yet the Vannevar Bush report, in many ways, was an exception to the rule for it was years ahead of its time in the performance and

effectiveness debate within the R&D community in general and private sector R&D in particular.

What follows below is a historical overview of the theoretical and practitioner literature as it relates to the notion of performance measurement within R&D. We have deliberately chosen a chronological order for presenting these ideas so as to recreate a sense of the debate that has waged within the R&D community over the last half-a-century. This approach provides a sense of the evolution as well as redundancies, repetitions, and weaknesses that have plagued this evolution. It also leads to certain conclusions about the manner in which this debate had played out and been influenced by concurrently on-going debates on performance measurement in other areas of human endeavor (e.g. education, healthcare, governance) At the end of the chapter, we bring this entire discussion together into an assessment of what we know and don't know about the R&D measurement problem.

2.2—Review of Literature on R&D Measurement (1959—present)

There has been no dearth of techniques and frameworks for measuring R&D performance. One author—during an extensive search of literature from 1956 to 1995—identified over 95 articles, 12 books, and a couple of research reports describing various techniques (Werner and Souder, 1997). The earliest documented reference regarding performance measurement aimed at the R&D function that we found was a 1959 article in Harvard Business Review by Melville H. Hodge Jr. titled "Rate Your Company's Research Productivity". It laments the poor state of data availability and analysis of R&D productivity by pointing out the fact that "while there is considerable subjective evidence that the industry is paying heed to…[the need for realizing the economic impact of fundamental research]…there has been little objective, quantitative data offered as to the amount and quality of basic research done by various industrial companies." (Hodge Jr., 1959)

The article identifies a number of difficulties in measuring performance of research programs in companies. It stresses upon the fact that while one should be able to measure productivity (and hence outputs and inputs of research,) often the absence of required data leads companies to measure inputs only. It suggests a number of methods and metrics, such as, commercial success, patents, and publications as ways to measure research output. The article ends with an analysis of research publications in five areas of physical and mathematical sciences but very carefully makes the point that, "this method must be used

with restraint, lest what it measures—scientific publication—replaces what it seeks to measure—scientific productivity—as a primary goal of the company's scientists".

Quinn (1961) questions the prevailing wisdom in taking research performance for granted and compares the state of development of research management at the time of its publication with that of marketing and manufacturing in 1920s. It calls for an end to the attempts of trying to find aggregated measures of performance that "do not exist" and instead stresses the need to look for measures that represent important determinants of research success. In one of the earliest—almost prophetic—statements in support of multi-dimensionality of research performance, it notes that "it is the structure of evaluation that is important…[and while]…even the most effective techniques available today…will surely be improved in the future…the structure for segmented evaluation of research output should stand the test of time."

Further elaborating on this segmented approach, Quinn (1961) identified three aspects of research evaluation, namely, measuring the economic value of technology (as apposed to cost of research,) the amount of technological output (regardless of economic impact,) and degree of congruence with corporate goals. In a thoughtful treatment of the issue, the article touches upon various important concepts and constructs in R&D performance measurement—such as, the appropriateness of various types of assessment techniques under different situations (e.g. rigorous qualitative assessments vs. pseudo-precise mathematics,) the real value of being able to assign numbers, the problems in using efficiency standards, the importance of incorporating quality, the place of "faith" in performance assessment, and the effective use of its results. These important philosophical issues, as we would argue later in this section, would remain lost, or at least would not be as well elucidated, elsewhere in this literature for many years to come.

Ranftl (1978) is a comprehensive study of R&D productivity commissioned at Hughes Aircraft Co. in 1973. The primary objective of the study was to "identify useful techniques for optimizing productivity in the R&D environment." Steering clear of definitional issues that plagued earlier and later studies, Ranftl (1978) used generic profiles of employees, managers, leaders, and organizations to define standards of productivity in R&D. It justifies this approach by asserting that "because quantitative measurement of R&D productivity is so difficult, many managers have developed, through their experience, informal guidelines or indicators for identifying productive individuals and organizations."

Ranftl (1980), however, concedes that "tasks that have been traditionally unmeasurable may, in time, become measurable, only to be replaced by more advanced unmeasurable tasks". It defines a productive organization as one that is effectively staffed and people-oriented, has high standards, operates in a sound competitive mode, has a creative and productive atmosphere, and a "can-do" attitude and esprit de corps among its employees. While many would agree to the above being desirable attributes, generally, they may disagree with the specific details, given the differences and diversity in the R&D landscape. According to this view, the search of an effective R&D performance (or productivity) measurement system is an elusive goal since the definitions of both R&D and effectiveness remain moving targets.

These differences notwithstanding, Ranftl's study made a significant contribution to the literature on R&D productivity and effectiveness. The study classified R&D productivity into three related factors, namely, personal factors, job-related factors, and project-related factors and consequently focused on improving operational, managerial, and employee productivity within R&D. The two phases of the study involved a highly participative process and depended on interviews of a wide cross-section of R&D managers and consultants—59 participating entities across a broad cross-section of organizations, a survey of over 2300 R&D managers, involvement of 28 prominent consultants, attendance at 23 productivity seminars, and an extensive review of the literature. Presenting its findings in a best-practices-like format, the study identified a host of productivity enhancing techniques and tricks, starting from personal productivity tips for R&D managers, to tips to improve employee morale, to tips about effective job assignment practices, to lists of effective planning practices and control techniques. The study also provided comprehensive checklists and templates for organizations seeking to adopt its recommendations.

Collier (1977) while referring to Ranftl study (1978) questions the wisdom in assessing productivity without relevance to fundamental goals of the organization. In essence, it was introducing the differentiation between efficiency (or productivity) and effectiveness. It differentiates between three stages of the process that begins at concept generation and ends at the realization of corporate goals (profits) from research and development.) It, however, maintains that these "steps are not simply additive, but utterly interdependent…[so that]…the productivity of the overall system is not the sum of individual productivities at each step, but their product". It further elucidates an imperfect yet evolving methodology to calculate R&D performance at each of these three stages as applied at Borg-Warner Corporation.

Mansfield (1981) focuses on economic payoffs as a measure of R&D's performance and success in the industry. He asserts that "to make effective use of R&D capacity, a company must first spell out its business objectives and then communicate them to its scientists and engineers. Research, after all, makes sense only when undertaken in areas relevant to economic goals". He talks about a number of different methodologies of making that assessment, namely, PROFILE (Programmed Functional Indices for Laboratory Evaluation)[10], QUEST (Quantitative Utility Estimates for Science and Technology), and PATTERN (Planning Assistance Through Technical Evaluation of Relevance Numbers).

Schainblatt (1982) is interested in productivity of scientific and engineering groups, programs, or other organizational units rather than individuals. A survey of 34 IRI member firms resulted in responses from 7 of them (20% of the sample) indicating that four firms regularly used performance or output indicators as a part of their review of R&D programs and three occasionally did so. While the sample is severely limited, the paper introduces a number of interesting concepts in performance measurement, namely, a return-on-research index, value of business opportunities generated, numerical indices of outputs, time spent on activities, resource allocation indices, and program value etc.

Packer (1983) aims at striking a balance between subjective (i.e. based on subjective assessments (e.g. "I think they are doing a good job",) surrogates (i.e., "since our budget is growing, we must be doing well",) and "structured" (i.e. based on objectively measurable criteria) approaches to measuring productivity of R&D. It also goes a step further by trying to bring insights from measurement theory used elsewhere to bear on the R&D performance measurement problem. The study also raised a host of very important issues like multi-dimensionality of performance, highly inter-correlated indicators, reliability, validity, verifiability, bias, relevance, goal congruence, psychological acceptability, and cost-effectiveness. The article presents a potentially useful technique of using factor analysis and mapping objectives-activities-indicators to arrive at an integrated measurement system that deals with many of the issues identified earlier.

Moser (1985) highlights the fact that commonly accepted measures of performance that are applied to R&D are not satisfactory and "provide no clues as to whether these results were obtained efficiently, or even if they are in fact accurate indicators of performance". It bypasses the question of trying to find appropriate criteria to measure performance and instead focuses on how performance is being measured in R&D units. Pappas and Remer

[10] PROFILE and QUEST were developed for the U.S. Navy while PATTERN was developed by Honeywell

(1985) present the now well-established idea of using different types of techniques (e.g. qualitative, semi-quantitative, and quantitative) for different stages of R&D (e.g. basic, exploratory, applied, development, and product improvement). They assert that before measurement of productivity is undertaken, top corporate managers must do two things, namely, decide what they expect to get from a research center, and determine the intent of the productivity measurement system. They then present a number of different techniques that can be applied at various stages of R&D.

Foster et al. (1985) presents an interesting model of R&D aimed to help think about the constituents of R&D effectiveness, such as, productivity (further sub-divided into potential productivity and technological development efficiency) and yield (further subdivided into potential yield and operating efficiency) and makes the case for taking steps to measure and manage each of these components separately. Several articles talk about using activity analysis and workload analysis to identify and remove waste (especially the use of engineers-scientists' time in non-creative administrative activities) and improve the creative productivity of the technical staff (Bolte, 1986 and Leifer and Burke, 1994).

Kinlaw (1986) describes a nominal group technique (NGT) based participative process to develop performance measures for Aerospace R&D. It identifies five different groups of performance ratios, namely, effectiveness, quantity, quality, value, and change. While acknowledging that this set of ratios is not unique, it highlights the need for further study of various types of measures and evaluation processes to derive an ideal and comprehensive set of measures. In the similar vein, Gronow (1987) reviews a number of methods of scientific evaluation—among them, peer review, interview and questionnaire, qualitative methods, and case studies--and their contexts. While concluding that no "accepted rationale [exists for] guiding the choice of method for a particular purpose or circumstance," it points out the need for empirical study of the effect of the different kinds of evaluations.

Brown and Sevenson (1988)—for the first time—introduce the concept of an R&D Lab as a system thus making a case for measuring different attributes, such as, quality, quantity and cost by looking at inputs, process, outputs, and outcomes of the system. They also highlight the fact that R&D managers should pay primary emphasis on outcomes rather than behaviors, though the former is more difficult to measure, and notes that only 20% of R&D managers even measured their productivity. They also stress the need to make measurement systems simpler, more objective, and focussed on few important measures.

Cordero (1990) explicitly puts forth the need to measure effectiveness and efficiency separately by measuring outputs and resources respectively. It also, mirroring once again the sentiment expressed in Collier (1977), talks about different types of performance (e.g. overall, technical, commercial) and the organizational level at with each should be measured (e.g. firm or SBU, R&D unit, and marketing/manufacturing unit.) The paper stresses an idea that many others apparently miss out i.e. technical performance should be measured by comparing technical inputs to technical outputs (e.g. R&D expenditure to number of patents) and commercial performance should be measured by comparing monetary measures of resources used to profits or cash flow generated. Underlying this is the notion that it takes much more than R&D's contribution to make a product commercially viable and, therefore, to hold R&D to account for profits generated might not be a fair performance indicator. Grady and Fincham (1991), however, misses this insight in subtle way when it talks about converting research from "functional cost center" to a "professional client service".

Robb (1991) takes a more comprehensive approach towards R&D performance assessment and presents four different approaches of productivity measurement that work at various levels ranging from intuitive-qualitative (e.g. answering a question: "how would things be had there been no R&D?") to extremely quantitative (e.g. counting outputs, transitions and dollars). The paper assets that "no one measurement suffices by itself…but these…taken together…tell a [positive] story of productivity and contribution to the company's performance". Thor (1991) likens R&D organizations with other white-collar organizations and explores the possibility of similarities between performance measurement systems of the two. It presents a "family of metrics" approach to measuring R&D organized in the form of an objectives matrix with performance targets identified. Using a scoring technique that combines this family of measures into an integrated index of performance, Thor (1991) provides a basis for using performance information for gain-sharing and incentives programs.

Szyonki (1990a, 1990b) present a vast number of tips—insights gained from experience of working with companies—on how to manage R&D effectively. Szyonki (1994a, 1994b), while building upon the earlier set of articles, laments the fact that everyone claiming to measure the effectiveness of R&D seems to be measuring outputs only. It provides a broad-based framework for measuring the effectiveness of R&D that talks about ten different activities (or functions) that R&D organizations must perform to be effective. These include, selection of R&D projects, planning and managing projects, generation of new product ideas, ensuring quality of R&D processes and methods, motivating technical people,

establishing cross-functional teams, coordinating R&D and marketing functions, transferring technology to manufacturing, fostering collaboration between R&D and finance, and linking R&D to business planning. This article provides a 6-point subjective assessment scale that measures the health or quality of each of these ten activities on the basis of presence or absence of certain systems, processes, or methods that might indicate strength or weakness in each area. In a relatively crude way, this article applies the concepts included in the advanced versions of the Balanced Scorecard's strategy maps to the R&D problem.

May and Pearson (1993) is a comprehensive review of total quality management (TQM) as applied to R&D enterprises. It begins by defining TQM's basic principles and reviews prior literature on TQM and R&D. The paper cites several prior studies whereby savings and performance improvements were achieved through the application of TQM within R&D settings, most notably, at Kodak Research Labs in US (Murray, 1985), ESSO Research Center and Exxon Chemicals Ltd. (Price and Gaskill, 1990), and ICI Chemicals and Polymers Ltd. (Hedge, 1990). The paper makes the point that TQM's principles are applicable to R&D, albeit with minor changes in the philosophy and language, so that "skeptics are not given ready ammunition". However, they acknowledged that "difficulty in measuring change makes implementing TQM an act of faith in R&D where it is not possible to place a precise financial value on the benefits". Of the 14 companies studied during this research, TQM was identified as "successful" by interviewees in all but one company. The judgment was, however, based on qualitative perceptions of success rather than a quantitative and verifiable metric.

McLaughlin (1995) is another comprehensive account of application of TQM in R&D, as is Uchimaru et al., (1993). The former presents a total quality model of R&D that emphasizes upon a balance between the three systems an organization is comprised of, namely, management, technical, and the social systems. It asserts the importance of measurement and assessment as a means to drive performance improvements and highlights the need for the performance criteria to be "multi-faceted (i.e. quality, productivity, cost-benefit, cycle time improvement,) interconnected and related to one-another".

The latter talks about the experiences of the authors—Japanese quality experts—in apply TQM to NEC Integrated Microcomputer Systems Inc. during 1980-1987 timeframe. It is a comprehensive study of installation of a performance enhancement system in a technical environment, rich in details on issues like engineers' resistance to measurement and management, organizational diagnosis, implementation steps and challenges, and results of

TQM implementation. On the latter count, the authors report improvements on virtually all measures of significance. Productivity of individuals and milestone achievement rates went up; design quality improved; and design learning curves and appreciation of customers latent needs improved. The company also received a Deming Prize in 1987.

Not all authors, however, are so positive and optimistic about TQM's track record in R&D settings. Schumman et al. (1995) assert that "R&D organizations have had only limited success so far from apply quality principles. Chatterji and Davidson (2001)—on the basis of decreasing publication of TQM related articles in R&D practitioner literature – ask whether that is a sign of decreasing interest and growing disillusionment within the R&D management community with TQM? "Did all the interest and activities in TQM come to nothing for the R&D management community?" they ask.

"R&D efforts to adapt and apply quality management principles have left some subtle yet powerful legacies—and as a community, we are all richer for these legacies", they assert. Among these legacies is the change in mindset of R&D managers concerning the potential for continuously improving management practices and processes, and greater receptiveness to ideas such as "what gets measured, get managed". They highlight the fact that TQM was a "company-wide galvanizing force" that taught R&D to learn and speak the same language as the rest of the company, and thus despite the apparently dying interest and lack of demonstrable benefits, TQM has played a constructive part in helping improve the performance of R&D.

Tipping et al. (1995)—in a curious admission almost four decades into the R&D measurement debate—again highlights the notion that an effective way to measure R&D performance is still elusive. "Even those world-class corporations that do look seriously to technology for competitive advantage are frustrated in their attempts to couple their R&D effectively into their businesses, in part because of the absence of an accepted methodology to measure effectiveness (value) and continually improve their R&D." Subsequently it presents the IRI's Technology Value Pyramid (TVP) model to measure the performance of R&D. Technology Value Pyramid is a modification of the Balanced Scorecard adapted to the R&D organizations.

The TVP represents a significant departure from earlier literature on R&D performance measurement—with the probable exception of Brown and Sevenson (1988)—that provides a hodgepodge of insights and techniques—sometimes very useful though--without really providing an underlying structure to them. Tipping et al. (1995) provide an organizing

structure that, at first look, appears like a comprehensive measurement system that adds value, in the words of the authors, "[by] integrating various aspects of previous work in a new framework for R&D and business management". The main focus of TVP, however, is on measuring effectiveness of R&D in the private sector.

Chester (1995) warns against the limitations, rigidity and manipulatability of performance measures by asserting that performance measurement is risky as it is often motivated and influenced (e.g. in terms of the factors measured and the weight given to each) by a company's past successes and failures and, in doing so, it runs the risk of being so optimized that it might not anticipate paradigm shifts, and in the words of the author, "miss the opportunities that lie outside the box". It introduces the element of incentives into the equation by highlighting the fact that performance measurement in and of itself does not guarantee motivation for improving productivity and must be linked to incentives.

Schumman et al. (1995) contend that "R&D is too complex a subject for a few measures to satisfy all needs…however, if R&D is viewed as a process, performance measurements can be effectively determined". It builds upon Brown and Sevenson (1988)'s process model of R&D and arrives at a six dimensional model of R&D with the dimensions being people, process, output, internal customers, external customers, and society. It then describes a matrix for selection metrics corresponding each of these six dimensions, to be used for performance tracking, in-process tracking, end-of-process tracking, and external tracking purposes. This gives rise to 24 cells of the matrix with several metrics placed within each of these cells. Rather than presenting this as generic matrix applicable to all R&D settings, the authors encourage R&D organizations to develop their own specific matrices that work best for them.

Miller (1995) asserts that "quality is another name for effectiveness in R&D". He presents a new framework to think about quality in R&D that departs from the earlier quality frameworks in commercial settings. This framework consists of four dimensions, namely, administrative and research processes, strategic decisions, cross-functional integration and project management, and participation of scientists. The study also identifies—on the basis of an analysis of 45 firms--four clusters of organizations on the basis of their mission and degree of reliance on quality practices in R&D. These clusters consist of firms managing at the scientific frontier, managing in revenue-dependency schemes, managing for TQM integration, and managing in strategic arena. It asserts that the findings are consistent with maturity models proposed elsewhere whereby organizations go through stages of evolution from lower level of integration of R&D to strategic alignment. The study concludes with the

assertion that "provided certain irritants are dealt with, the quality approach is not only credible in R&D community but is a valuable addition to efficient R&D management".

Davidson and Pruden (1996) present another quality perspective on R&D management through a four-stage model of a "Total Quality Transformation Grid for R&D". They define the four stages of performance measurement through which an R&D organization goes through with the first stage characterizing ill-defined value-added and management by exception; the second includes customer satisfaction and forms of product and process variability assessment; the third using a continuous process improvement approach in addition to its predecessor; and the fourth employing a multi-dimensional performance measurement to optimize processes.

Hauser and Zettelmeyer (1996) presents another framework to look at R&D metrics and measurement. It classifies R&D activities into three tiers, namely, tier-1 activities (those that form organizational objectives or "exploratory" activities e.g. basic research,) tier-2 activities (those that select and develop "programs" to match and create organizational core-competencies,) and tier-3 activities (those activities or "projects" focused on more immediate needs of the customer.) It asserts that that the choice of performance metrics should be driven by the type of R&D activity being performed—as described by the three tiers of R&D activities—and metrics suitable for one type of activities may be counterproductive for other type of activities.

NRC (1996, 1997) presents a multi-dimensional model of performance for R&D organizations. Seeking to identify, define, and understand the various dimensions of "world class" performance in research, development, and engineering settings, it defines five pillars, namely, customer focus, resources and capabilities, strategic vision, value creation, and quality focus. It identifies 25 characteristics within the five pillars of performance. It uses four qualitative descriptors (i.e. poor, adequate, good, and excellent) for each of the 25 characteristics identified above. While this, like Tipping et al. (1995,) is another comprehensive attempt to introduce performance multi-dimensionality in R&D, it falls short of replicating the Balanced Scorecard (Kaplan and Norton, 1992, 1993, 1996, 2001). For example, three of the five pillars (i.e. customer, quality, and value-creation) of world-class R&D organizations deal with the customer perspective of the Balanced Scorecard. The financial and internal process perspectives are combined in one resources and capabilities pillar—thus allowing the possibility that one or both could be shortchanged in the course of measurement.

Kostoff (1997, 1999, 2000 etc.) and several of the other publications by the same author spanning almost the entire decades of the 1990s is an interesting collection of work on numerous individual pieces of the R&D metrics and the measurement problem such as, peer review, retrospective analysis, quantitative metrics, cost and benefit analysis, citation analysis, technical roadmaps, impact assessments, and integrated metrics. The overarching contribution of this line of scholarship is an in-depth treatment of each of the various type of metrics used in R&D evaluation—facilitated in part by the author's vast experience of being engaged in the R&D evaluation enterprise in the public sector. No other reference, we found, provided such an extensive review of issues and intricacies of such a varied mix of evaluation schemes and metric types.

What Kostoff (1997, 1999, 2000 etc.) and others do not do as well, however, is to develop a system. In a similar vein, Ellis (1997) is a comprehensive piece of work on a number of disparate issues in evaluation of R&D processes, such as, process measures, financial measures, quantitative measures, assessing the health of technology transition and R&D's interactions with other functions, internal and external evaluations, and leading and lagging indicators etc. While it aggregates a wealth of information from a vast number of sources into a comprehensive reference on the minutiae of R&D evaluation, it does not provide an organizing framework to put everything together into a measurement system.

I.C. Kerssens-van Drongelen (1997) builds upon the measurement system thinking—first subtly eluded to in Brown and Sevenson (1988) and Tipping et al. (1995)—and introduces a number of concepts important for thinking about such systems. It presents the need for performance measurement as an essential component of quality management in R&D. The article ties the R&D performance measurement system with a well-established literature in engineering and management, namely, the theory of control systems design. Among the new concepts introduced in the article are a more explicit discussion of functions of measurement (e.g. motivational and diagnostic,) a more detailed treatment of effect of contingency factors on measurement system features, and a better-developed list of major system design parameters (e.g. metrics, structure, standards, format, and frequency.) The study also introduces the idea of Balanced Scorecard as a way to deal with the problem of performance multi-dimensionality.

Kerssens-van Drongelen et al. (1999) builds upon the model presented above and provides an empirical basis to the earlier study. Kerssens-van Drongelen (1999, 2001) are important contributions to the literature and a comprehensive reference on one approach to thinking about an R&D measurement system. Among their many contributions are: an in-depth

review of earlier literature, a taxonomy of various variables and contingency factors that might affect measurement system design, the design of a comprehensive R&D measurement system, and theoretical and empirical insights on how might it look like for different states of the taxonomy. We would extensively use and build upon this framework during the course of this work.

Geisler (2000) is a fairly comprehensive reference on science and technology metrics used in a variety of contexts, such as, inputs, outputs, and process metrics, on the one hand, to financial, bibliometrics, patents, and peer reviews etc., on the other. In addition, it presents applications of evaluation schemes derived from these metrics in areas like academic science, industrial innovation, and public sector science and technology. Pearson et al. (2000) touches upon the issue of measurement in R&D and presents as the growing consensus the feeling that "return on R&D expenditure cannot escape close scrutiny, notwithstanding the well-recognized difficulties of evaluating the benefits of R&D". Building upon the Kerssens-van Drongelen (1997), it also stresses importance and popularity of broader-based techniques such as Performance Pyramid (Lynch and Cross, 1991), Balanced Scorecard (Kaplan and Norton, 1992), and Intellectual Capital Model (Edvinsson and Malone, 1997). It also introduces the idea that these varied techniques are complimentary rather than mutually exclusive and hence best used under circumstances appropriate for their application.

Baglieri et al. (2001) essentially re-packages some of the ideas expressed earlier, namely, stages of R&D (reworded as "domain of measurement" e.g. generation, transition, diffusion) and types of metrics (reworded as "object of measurement" e.g. input, output and process) but presents them in the form of an integrated measurement framework that rests on the additional insight that each of the stages of R&D (also sometimes referred to as basic, applied and development) can have its own unique set of input, process, output and outcome metrics. It further elaborates the framework to include the idea of knowledge on the shelf (Kos) and knowledge in progress (Kip) and asserts that the value-added of the integrated approach comes from being able to apply different methodologies suitable to measure each of the above components while operating within the same general framework.

Jordan and Malone (2002) addresses the issues faced by public-sector R&D agencies in the light of GPRA and OMB-PART in the 1990s. It highlights the importance of establishing a framework for doing assessment of R&D programs with the expectation that "the elements and flow of the framework would depend upon the type and purpose of research". It

introduces program logic models as one such candidate framework. It asserts that while many organizations may not use the terminology, they in fact use a logic model—comprising inputs, activities, outputs, and outcomes—as a basis of their assessment frameworks. The paper covers a lot of ground in reviewing well-established tools and techniques, e.g. peer review, economic indicators, case studies, bibliometrics, and benchmarking.

Jordan et al., (2003) uses a Competing Values Model to structure attributes of organizational excellence along several dimensions. Through several focus groups at actual R&D laboratories, the authors arrive at thirty-six attributes (distributed in four categories) of organizations that foster excellence in research. The four categories include human and physical resource development, innovation and cross-fertilization of ideas, management and internal processes, and setting and achieving relevant goals. The score on these thirty-six attributes provides a profile of an organization's research environment rather than a single indicator of its research excellence. Based on experimental implementation across several labs of the Department of Energy, the authors find the resultant information quite useful in identifying weak areas in organizational management and for insights about performance enhancement.

Loch and Tapper (2002) is a specific application of Kaplan and Norton's Balanced Scorecard to an applied research group in a geological research setting. It carefully formalizes several process dimensions of the Balanced Scorecard—a fact that several other authors either miss out or only pay lip-service to (including Kerssens-van Drongelen, 1997 and 1999)—like cascading and operationalization of strategy, alignment with organizational strategy, employee empowerment, performance multi-dimensionality, and scorecard-development etc. Similar to the performance perspectives of the Balanced Scorecard (Kaplan and Norton, 1992, 1993, 1996, 2001) it develops a strategy-driven performance scorecard for the organization in question. It identifies several benefits of the scorecard development exercise and the scorecard itself, namely, employee empowerment, bottom-up and top-down integration, project prioritization, and benchmarking. Finally, it finds several positive effects of putting the system in place. For example, it reports dramatic improvement in actual vs. forecasted performance on budget, improved innovativeness and idea generation, a tripling of quality-adjusted publication activity, improved transparency about projects and accomplishments, and improved focus on outcomes that mattered among the researchers and scientists.

To sum up, notwithstanding the gradual improvement in the state of the art over the time horizon covered, there are significant deficiencies and inconsistencies in the way the field of R&D performance measurement has evolved. For example, many of the more profound insights that were expressed as back as 1960s (e.g. Quinn's 1961 article in HBR) were lost and then rediscovered years later. There was also relatively limited cross-germination of ideas from performance measurement literature that was simultaneously evolving in other areas, such as, education, governance, and health (e.g. only Packer (1983) attempts to consciously make that connection) as well as across the different sub-communities (e.g. the academic , business, and practitioner communities as evidenced by publishing in different journals i.e. mainstream business journals vs. professional journals vs. trade and technical journals) active within the same broad area of knowledge. Similarly, as late as 1990s, there was little talk of frameworks and systems of measurement and the literature mostly focused on individual methodologies and techniques presented in a hodgepodge manner.

Another feeling that one gets after looking at this literature is the apparent lack of methodological sophistication. For example, the R&D measurement literature does not interface well with that on measurement theory and is astonishingly naïve about issues like measurement error, multi-dimensionality, correlations, causality, reliability, validity, variability, and bias etc. Some exceptions are Quinn (1961) and Packer (1983). Also, while there is certainly some progress over decades, one feels a curious lack of progress on shorter timeframes in that similar unresolved issues are being repeated over and over again and as if the literature is circling around familiar issues and not finding resolution. The notion of an R&D system, for example, did not take root before Brown and Sevenson (1988) and the system view of performance measurement was not thoroughly and meticulously investigated prior to Kerssens-van Drongelen (1999). The latter was the first systematic attempt to use systems control theory to arrive at what a control system must entail and how to use that guidance to develop a performance measurement system.

Finally, even after four decades of development, there are still voices of dissent and dissatisfaction from the state of the art. For example, as late as 1990s, Bukowitz and Petrash (1997) talk about a "lack even [of] a basic language for talking about intellectual [assets]." Earlier, Francis (1992) asserted that "as in all other operations…R&D must be subject to effectiveness measurements, or payoff assessments". Similarly, Jamsen et al. (2000) presents some empirical evidence suggesting a significant mismatch between the expressed needs and the actual practice of performance measurement in R&D. Thor (1991) also questions the state-of-the-art in performance measurement in R&D by asking: "conceptually, research organizations should be much like other white collar organizations. Why, then, there has

been little progress in developing performance measurement indicators for those types of firms?"

While one logical explanation for the repeated expressions of frustration is that each new significant addition to this literature gradually "raises the bar" on what one expects from a well-functioning performance measurement system, there is some element of truth to this frustration too. While a careful glance at the literature does present some evidence of growing sophistication (e.g. a systems model of R&D presented by Brown and Sevenson (1998) ends up becoming a de-facto foundation for many other models of R&D performance, like, Kerssens-van Drongelen (1999) and Baglieri (2001) etc.) there is still quite a bit amiss in the literature. The connections are far too few and haphazard to depict a field that is reaching a resolution on at-least some basic issues and moving on towards issues of higher sophistication, as has been a case with other areas e.g. healthcare (Donabedian, 1980) and governance (Kauffman et al., 1999, 2001 etc.) Also affecting this lack of resolution in the literature is the noticeably absent statistical evidence on actual efficacy of performance measurement techniques and systems, above and beyond some anecdotal evidence (May and Pearson, 1993) and descriptive and largely unverified organizational case studies (Patterson, 1983; Uchimaru, 1993; Krogh et al., 1988; Foster, 1996). One possible exception is Loch and Tapper (2003) that carefully documents the effects of implementing a performance measurement system in a research setting.

In all, the picture presented through the above review of theoretical and practitioner literature is that of a nascent yet fast maturing field of intellectual discourse. It also represents a field that is growing in its sophistication of the treatment of various issues that it is concerned with—philosophical, pragmatic, and methodological.

2.3—Perspectives from Performance Measurement Literatures in Governance, Education & Health

As we talk about the deficiencies in the R&D measurement literature, it would be useful to briefly contrast that with comparable literature in some other areas, namely, governance, education, and healthcare. This could be helpful in determining not only how the problems faced by these distinct intellectual streams might be similar with R&D but also how they might be different and why. Surely, many of the problems that have plagued the R&D measurement debate have also preoccupied those who have searched for similar answers elsewhere. For example, the notion of what constitutes the most important dimensions of

performance has been an open question—settled only, if ever, through years of debate within the respective communities—in all three literatures that we looked at.

At the first instance, we find considerable debate and controversy on definitional issues. For example, what constitutes good governance? is a question far from settled in the literature. Multiple definitions of the phenomenon exist (e.g., Frischtak (1994), Ear (2002), Bratton and van de Walle (1997), and Dethier (1999),) but also a healthy debate ensues on whether governance is a simple phenomenon or an integrated mix of several related phenomena? Similar discussions can be found in the healthcare (i.e. "What constitutes good quality care?") and education (i.e. "What constitutes quality education?") literatures. Researchers have even wondered (and continue to do so) about what is the "core process" of education delivery and how to map, benchmark, and improve it (Stecher, 2005). Clearly, these discussions are not very dissimilar from the ones that go on within the R&D management community which too struggles with questions like: what constitutes "good" R&D management practice; what does "quality" mean in R&D settings; and whether or not there is a "process model" of R&D.

Moving forward from the philosophical-definitional concerns, toward more pragmatic-measurement issues, the governance literature struggles with trying to identify quantifiable dimensions of the phenomenon. For example, we find a healthy debate on what constitute the dimensions of good governance. The maximalists and the minimalists argue over what should be the characteristic of the phenomenon and what should be its effect (Munck et al., 2000). The most important contribution in this regard comes from Kaufmann et al. (1998, 2001 etc.) that uses cross-country data on 174 different measures from 17 different sources to construct quantitative indices for six dimensions of governance, namely, institutional quality, government effectiveness, rule of law, control of corruption, political stability and regulatory quality. There is considerable debate and disagreement, however, over whether or not these six dimensions sufficiently reflect good governance with other authors arguing for including constructs like lower inequality, greater diversity, maintenance of traditions, high quality bureaucracy, successful provision of public goods, effective spending, democracy and political rights as precursors and constituents to good governance.

Similarly, it is widely accepted that healthcare quality is a multidimensional construct, and multiple indicators, each providing insights into the different domains of quality, are needed to evaluate it (Rubin et al., 2001) More often than not, what ultimately becomes the gold standard of measurement has little to do with actual completeness and validity in capturing the underlying phenomenon and more to do with availability of data and

charisma and perseverance of its champions. There are several examples, and some exceptions. In the governance literature, for example, Munck et al. (2000) provides a fairly detailed analysis of the biases associated with well-known indices of democracy and freedom (e.g. Freedom House Index and the Transparency International's Corruption Index) that cast considerable doubt over the ability of these to measure the underlying constructs of political freedom and corruption respectively, yet these remain well-accepted and widely used in their relevant communities.

Similar concerns are raised in the healthcare and education literatures—one instance of which is the assessment of quality of health outcomes of individuals through a standard known as SF-36 (short for "Short-Form 36"). SF-36 is a health status questionnaire comprising 36 items that are asked from patients to assess their health care quality. Developed and championed by Dr. John E. Ware, SF-36 has, despite early criticism and critique, become a de-facto standard in healthcare community with even shorter versions of its—SF-12 and SF-8—being developed and acknowledged as useful indices of quality of life of an individual.

One interesting idea that has been experimented within the healthcare and education literatures—but not so much in the R&D performance literature—is that of evaluation from the perspective of multiple stakeholders. An interesting vignette coming from quality of healthcare literature is the apparent disconnect between the physicians' and patients' perception of what constitutes quality of care. While the former believe that the most important element of care is technical excellence (i.e. doing the right diagnosis or procedure every time) patient surveys consistently show that they cared more about interpersonal interactions and factors like waiting on the phone or in a waiting room before getting to the doctor (Brook et al., 2000). Needless to say, these surveys have resulted in the development of evaluation schemes that incorporate perspectives of multiple stakeholders, namely, objective and subjective measures of health quality gleaned through expert opinions (physicians' perspective), patient satisfaction surveys (patients' perspective), and providers' perspective. A discussion along these lines in the R&D management community where scientists' perspective on what is interesting and useful science sometimes differs from those of their financiers (e.g. corporate top-management, or a public funding agency) might potentially lead to some fruitful insights.

Technical issues like measurement error, bias, validity, and reliability etc. that have escaped an in-depth discussion within the R&D community have long been recognized, debated, and incorporated in education, healthcare and governance literatures. Similarly, in the

healthcare literature, the notion of a healthcare "system" was first introduced by Avedis Donabedian (Donabedian, 1980) about a decade earlier than Brown and Sevenson (1988) suggested a systems model of R&D. Donabedian's (1980) systems model of healthcare required assessment of structure (inputs and organization), process (content and configuration) and outcomes (procedural end-point and impact)—a scheme much similar to Brown and Sevensons' input, process, output, outcome quartet.

Not unlike the unpredictability of R&D outcomes, the quality of healthcare literature has also struggled with the notion of unpredictability and uncertainty of healthcare outcome (McNeil, 2001.) The resultant consensus points toward a greater reliance on process measures (though not to the complete exclusion of structure and outcomes measures) despite the fact that they are less precise and more expensive to develop (Donabedian, 1980). Several steps need to be taken into account to systematically develop process measures for quality of care (Rubin et al., 2001). Consequently, the Evidence-based medical practice guidelines have developed to help standardize the process of care. Similar questions arise in the education literature with potential for developing a process-based quality improvement system (Stecher, 2004{}). This dynamic is also mirrored in scientific management literature—although at a much nascent level, qualitatively—as it calls for better adherence of best practices and the scientific method within the research community (Gauch, 2003).

The literature on governance, education and healthcare also inform about the adoption dynamics of particular measurement methodologies. These depend on the norms and traditions (e.g. perceptions of accountability, tendencies to entertain internal debates about efficacy and effectiveness, and the use of hard evidence and data to self-reflect, and governance structures) of a particular professional community. They also, at times, depend upon temporary or permanent political considerations.

Gormley and Weimer (1999) ably illustrate this point with an example that compares the adoption of organizational performance report cards in the healthcare and education sectors. They argue that for a variety of political reasons, the practice of using performance scorecards—when first introduced in the 1980s—was widely adopted in education while forcefully resisted and then altogether abandoned in healthcare. They use the analogy of *Downsian mobilization* as against *Schattschneider mobilization* to describe the adoption dynamics in education and healthcare respectively, whereby in the former the advocates for a particular solution mobilize to create new institutional arrangements to support it while

in the latter critics of a solution expand the scope of conflict followed by the destruction of the policy sub-system.

These two streams of literature also point out other peculiarities in adoption that might be useful to look at and learn from as we try to think about performance measurement systems in R&D. For example, there is considerable anecdotal and survey evidence to suggest that while lot of performance information on healthcare providers exists in public domain—e.g. since 1989 the states of New York and Philadelphia release data on results of coronary arterial bypass graft surgery performed at various hospitals allowing consumers to rate hospitals based upon their level of risk adjusted mortality rates (McGlynn and Brook, 2001)—people tend to disregard this information in making high-stakes decisions about choice of providers in healthcare (Marshall, Shekelle et al., 2000). This is not so in education where parents seem to be much more informed about school choices for their kids and make conscious decisions to relocate to neighborhoods with better quality school districts. The underlying causes for this differential use of critical quality and performance information is a not yet convincingly explained in the literature.

The provision of information on healthcare quality has grown in the US and with it the available array of "report cards", "provider profiles", and "consumer reports" (Marshall, Shekelle et al., 2000). Vaiana and McGlynn (2002) is a comprehensive of review of lessons that may be learnt from cognitive science about how to design consumer report cards and may be applied to healthcare, education, governance, and organizational report cards. The ability of these hospital (or provider) scorecards to discriminate between performance and processes of two hospitals (Krumholz, 2002) or their ability to provide meaningful and consistent information is under considerable scrutiny. Gandhi et al. (2002) and Simon and Monroe (2001) are skeptical of the tendency of the physician scorecards to provide inconsistent information. Similar inputs and outcomes-based scorecards have also propped up in the area of education (Russell, 2004). Both these efforts to provide accountability and influence performance through greater provision of information have encountered similar problems and have had similar effects (Stecher, 2004).

In addition to the above, there is also a relatively thin but growing literature on evaluation of performance improvement initiatives within education and healthcare (e.g. Lammer et al., 1996; Feeney and Zairi, 1996) as well as cross-fertilization of ideas and best practices across these literatures (e.g. Garside, 1998, Stecher, 2004; Vaiana and McGlynn, 2002). There is certainly a need for further dissemination of the evidence across various practice areas (e.g. education, healthcare, governance, and others) and to learn from the experiences of

individuals across these areas. The similarity of issues and problems at various stages of development within these literatures in striking. McGlynn (1997) enumerates six challenges in measuring the quality of healthcare within the United States and one can spot several commonalties with similar debates within education, governance, and R&D management literatures.

These and many more factoids and lessons must be kept in mind as we embark upon an effort to develop a system of performance measurement for research and development organizations as not only do they help us understand what might and might not work, but also help us anticipate and plan for bottlenecks and challenges that might arise throughout the process. Before we conclude this review of literature, however, it would be useful to consider how notions of performance and measurement have evolved in public sector R&D. Next, we look at that.

2.4—THE NEW FOUND IMPETUS FOR PERFORMANCE MEASUREMENT IN PUBLIC R&D

Evolving alongside the debate on performance measurement in private sector R&D is a similar debate among the practitioners of a large public sector R&D community responsible for conducting, funding, and evaluating public research programs. While the concepts of performance and measurement have evolved quite noticeably in the private sector all through the second half of the twentieth century, those in the public sector underwent a slower change. A different set of factors affect the conduct of public-sector research. For much of the half century, public sector priorities and decision calculus in the United States were dictated by the cold war.

In fact the roots of the very consensus that lasted for about four decades in this community can be traced back to the end of the Second World War itself. One statement in the Vannevar Bush report "Science: The Endless Frontier" that was the foundation for the public sector's views on issues of research sponsorship and performance signifies this consensus much more than any other and is worth repeating here. In the concluding chapter of the report Dr. Bush writes: "We submit, however, that the nation's need for more and better scientific research is such that the risk [of lack of cost-effectiveness] must be accepted."

Thus, throughout the cold war, the public sector considered the performance dimensions of efficiency and cost-effectiveness as secondary to effectiveness and robustness (and perhaps accountability), not only in the defense and national security realms but also in blue-sky

R&D in general. This, however, began to change in the early 1990s after the fall of the Soviet Union, and with it came a renewed pressure on the leadership of public sector R&D enterprise to rethink its performance measurement paradigm. An OECD study finds that three concerns are noticeable in the performance management of all OECD countries, albeit to different degrees: public concern that governments: a) improve performance, b) clarify responsibilities and control, and c) realize cost savings (OECD, 1997). Irwin Feller (2002) cites three other reasons for the focus on performance management in research programs: a) a need for accountability in times of flat budgets; b) concern that public funding is being provided in areas that do not obviously contribute to national well-being e.g. civilian technologies; and c) eroding political acceptance of claims by the research community, particularly by academics, of self-policing.

Two significant developments that amply reflect the public sector's catching up on the concept of performance measurement in R&D are worth noting here:

2.4.1—Government Performance and Results Act (GPRA)—A befitting culmination of the above mentioned chain of events in the public sector is the Government Performance and Results Act of 1993. The Act shifts the focus of measurement paradigm for public programs from inputs to performance and results (Cozzens, 1995.) The basic objective of GPRA-mandated reporting is to come up with measures of outcomes that can be tied to annual budget allocations. More specifically, the law requires each agency to produce three documents: a strategic plan, which would set general goals and objectives over a minimal 5-year period; a performance plan, which would translate the goals of the strategic plan into annual targets; and a performance report, which would demonstrate whether the targets were met (NAP, 1999).

GPRA provides a strong impetus for R&D evaluation by mandating federal agencies to set strategic goals and performance measures for managing and budgeting as a means to improving efficiency, effectiveness and accountability for all federal programs and spending. Research and Development is not excluded. In the context of government R&D, the purpose of this exercise would be to improve efficiency and effectiveness of public dollars spent on research programs (NAP, 1999). The GPRA especially stresses the use of metrics as it requires the annual performance plan and goals to be expressed (generally) in an "objective, quantifiable and measurable" form through performance indicators that measure or assess the relevant outputs, service levels and outcomes of each program activity (Cozzens, 1995.)

It also specifically differentiates, through a common vocabulary, between initial inputs, short-term outputs, and long-term outcomes. Since the passage of GPRA in 1993, agencies

have made increased efforts to come up with quantifiable, useful, and defendable measures and metrics for their investments. An inter-agency group was established to help agencies develop performance measures for research programs by initially implementing pilots that included programs from 20 agencies spreading across the entire spectrum of government functions and activities (Cozzens, 1995).

Struggling to meet GPRA requirements, many federal agencies funding or conducting in-house R&D have found themselves in a relatively uncharted territory, often struggling with very basic issues like how to measure the impact or performance of knowledge generating basic research? etc. Also there is a concern that too much emphasis on a measurement culture might in fact create wrong incentives for the people who are affected by them or encourage agencies to measure what is easy and neglect what is important (Cozzens, 1995.) Keeping in view these intricacies, GPRA also allows use of an "alternate forms [of evaluation]", as approved by OMB, for agencies that find it infeasible to express performance in quantitative form. Many agencies have experimented with these alternate forms.

A 1999 COSEPUP study calls for the use of expert reviews to evaluate federal research programs on three criteria, namely, quality, relevance and leadership. These expert reviews are designed to be much more than traditional scholarly peer evaluations and include inputs from users of federal research as well (NAP, 1999). A follow up for the 1999 study that documented agencies' efforts to comply with GPRA requirements validated that expert reviews—and not quantitative metrics—were indeed the predominant evaluation methodology among agencies that dealt with R&D (NAP, 2000). This study—among its many recommendations—called upon the agencies and their oversight bodies to "work together, as needed, to facilitate agencies integrating their GPRA requirements with their internal planning, budgeting and reporting processes."

2.4.2—Program Assessment and Ratings Tool (PART)—The Program Assessment and Ratings Tool (PART) is an analytical framework put forward by the Office of Management of Budget (OMB) to try to harmonize the assessment of the seven different types of federal government programs (e.g. competitive grants programs, capital assets and service acquisitions programs, regulatory programs, credit programs)—not just the R&D programs. The PART is a diagnostic tool—designed as a series of questions—that relies on objective data to inform evidence-based judgments to assess and evaluate programs across a wide range of issues related to performance. The questions are written in a Yes/No format and require the user to provide a brief narrative explanation, including any relevant evidence to substantiate the answer. This forces individuals filling out the evaluations to think about

the questions and formulate thoughts before answering them. The accompanying documentation clearly defines the purpose of the question, the standards of performance for a "Yes" or "No" answer on each type of question, as well as examples of evidence that might support each.

These Yes/No answers are then converted into quantitative scores through a simple weighting logic. This would result in the standardization of the performance evaluation process across all federal programs. The PART forms filled out for every federal program would then be available for the public scrutiny and review. PART was launched in FY2002 with a series of pilots and was expected to be fully deployed across all federal program and project portfolios in FY2004. OMB expects to use this tool for making budgetary allocation decisions across federal agencies and programs.

The PART guidance documents starts out by stating that: "the key to assessing program effectiveness is measuring the right things. The PART requires OMB and agencies to choose performance measures that meaningfully reflect the mission of the program, not merely ones for which there are data. The measures should reflect a sense of program priorities and therefore will likely be few in number. As a general approach, we expect these measures to reflect desired outcomes; however, there may be instances where a more narrow approach is more appropriate and output measures are preferable." (OMB, 2002)

The tool operationalizes this intent by dividing each PART evaluation into four sub-sections, namely, program purpose and design, strategic planning, program management, and program results. Each of the seven types of federal programs has a different PART framework specifically tailored to the needs of the type of program in question. In the PART for R&D Programs, for example, there are 33 questions in all (7 for purpose and design, 9 for strategic planning, 11 for program management, and 6 for results) in 4 sub-sections. Each question is equally weighted within its respective sub-section. The total score for the program is calculated by adding up individual scores on the 4 sub-sections and weighted in a manner that program results gets the highest weight of 50%, followed by 20% each for purpose and design and program management, and 10% for strategic planning.

Our conversations with an observer of the process that resulted in the formulation of the PART methodology suggested that these were developed quite independently of the GPRA process and were not motivated by GPRA itself. OMB's own guidelines, however, suggest that while GPRA might be a starting point for federal agencies to become more performance-driven, PART is really where the future lies—an assessment that might materialize if it becomes an integral part of the federal government's budgetary process. To

that effect, OMB's guidance document states *that* "GPRA plans should be revised to include any new performance measures used in the PART framework, and unnecessary measures should be deleted from GPRA plans."

Whatever might be the ultimate fate of GPRA and PART, whether they become just another "bureaucratic" requirement that agency managers would have to fill anyway or really encourage a performance mentality in the public sector, the trend towards greater appreciation of performance and performance measurement in public sector is quite unambiguous. Performance measurement in R&D might thus have finally come of age, not only as a need but also as a requirement.

CHAPTER—3

THE R&D BALANCED SCORECARD: ESTABLISHING A REFERENCE MODEL

The notion that measurement of an organization's (or a system's) performance must incorporate, to the extent possible, all key dimensions has been discussed in a number of literatures, including, education, healthcare, governance, management, and measurement theory itself. In education, for example, the fundamental question of what constitutes intelligence, how to measure it, and whether or not intelligence itself is a multi-dimensional concept has been widely debated (Gardner, 1993)[11]. In healthcare and governance as well the fundamental issue at stake is the very definition of quality healthcare (McGlynn, 1997)[12] and good governance (Kauffman et al, 2001.)[13]

In the business, management, and organizational contexts as well, the problem of performance multi-dimensionality is pervasive e.g. financial vs. non-financial performance, long-term vs. short term performance, and performance from the perspective of an organization's shareholders vs. performance from the perspective of its managers or employees. The earliest Balanced Scorecard tried to address one of the several manifestations of performance multi-dimensionality in business organizations, namely, the over-emphasis on financial metrics to measure an organization's performance and the consequent lack of appropriate measures in an organization's financial statement(s) to account for intangibles like customer satisfaction, innovativeness, employee morale, goodwill and brand recognition etc. (Kaplan and Norton, 1992 and 2001)[14] In this chapter

[11] The debate is informed by Howard Gardner's theory on multiple intelligences—seven in all.
[12] Healthcare researchers have found at least two major (and probably several minor) dimensions of quality of care, namely, technical dimension and the human dimension of care (McGlynn, 2001)
[13] The nascent but fast growing literature on good governance is still debating on the definition and various dimensions of good governance. One of the relatively better established definitions (Kaufmann et al. 2001) uses six dimensions to measure governance across countries.
[14] The Balanced Scorecard, in the words of Kaplan and Norton, was "first developed… in the early 1990s to solve a measurement problem. In knowledge-based competition, the ability of the organizations to develop, nurture, and mobilize their intangible assets was critical for success. But financial measurements could not capture the value creating activities from an organization's intangible assets: the skills and competencies, and motivation of employees; databases and information technology; efficient and responsive operating processes; innovation in products and services; customer loyalty and relationships; and political, regulatory, and societal approval. We

we review the intellectual development of the Balanced Scorecard in an attempt to derive a generic reference model for future analysis.

This chapter is organized as follows. We start with discussing a fundamental caveat in defining the scope of the Balanced Scorecard itself, in general, and this research, in particular, namely, the problem of establishing a reference model for the Balanced Scorecard. Section 3.1 describes the origins and motivations behind the Balanced Scorecard in somewhat detail and arrives at the basic definition of the Balanced Scorecard. Section 3.2 presents an analytical framework that encapsulates the various structural features of the Balanced Scorecard. We use this framework as a benchmark to compare other formulations and implementations of the Balanced Scorecards. Section 3.3 presents some of the criticisms and alternatives of the Balanced Scorecard and our own judgment of whether or not these measure up to valid criticism of the traditional Balanced Scorecard.

Section 3.4 presents a discussion and analysis of the strengths and shortcomings of some attempts to implement a Balanced Scorecard in R&D organizations. Section 3.5 presents our own generic R&D Balanced Scorecard as reflecting our understanding of what would constitute a useful variant of the original Balanced Scorecard adapted to the R&D settings. We close this chapter with a brief discussion on an analytical agenda to test the appropriateness of our generic R&D Balanced Scorecard. We subsequently build upon this analytic agenda to arrive at specific hypotheses in the methodology chapter (chapter-4) of this study.

Before we define what a Balanced Scorecard is, and lay out its structural and implementation details, a caveat is in order. Since its first launch over a decade ago, the term Balanced Scorecard has come to be used in several contexts and forms. This is almost an inevitable outcome when any idea enjoys wide popularity in the managerial and scholarly community (Hackman and Wageman, 1995.) Barring the basic insight of using multiple dimensions of performance that all these various types of multi-dimensional performance measurement frameworks (that have commonly and generically come to be referred to as "Balanced Scorecards") have in common, they differ in several structural and implementation details. According to a survey conducted by the Balanced Scorecard Collaborative—a consulting organization created by the initiators of the Balanced

---

proposed the Balanced Scorecard as the solution to this performance measurement problem" (Kaplan and Norton, 2001).

Scorecard—about 50% of all Fortune 500 companies in the United States claim to implement a Balanced Scorecard, but only half of these companies do so in its correct shape and form[15].

There are several possible explanations for the above discrepancy. In many instances, the consulting industry has caught on to the faddish popularity and appeal of the Balanced Scorecard and has started delivering solutions that are either a "repackaged" version of other performance measurement and management approaches (e.g. Six Sigma Balanced Scorecard of Gupta, 2004,) or are a close approximation of the Balanced Scorecard (e.g. Olve et al., 1999; Brown, 1996; Nevin, 2003; I.C.K. van-Drongelen, 1999 .) While some have so named these scorecards to reflect the intellectual debt they owe to the original Balanced Scorecard, others have attempted to highlight the differences. Regardless of the particular kind of association, however, these approximations often differ from the original Balanced Scorecard in subtle but significant ways.

Another reason for the diversity is that the Balanced Scorecard is itself a "moving target" that is evolving over time. Even if we disregard the effects of the herd behavior within the consulting industry, the fact that Balanced Scorecard was different a decade ago from what it is today (Kaplan and Norton, 1992, 1993, 1996, 2001) can itself be a significant source of the variation—primarily attributable to learning lags—introduced in the structure and implementation of the Balanced Scorecard over the years and across organizations. Finally, the Balanced Scorecard is, by design, flexible enough to accommodate the unique features of a diverse set of organizational forms (e.g. corporate sector, public sector, education, healthcare, defense)—a fact that can, in principle, introduce differences in structure and implementation. For all of the reasons cited above, we find considerable variation in the Balanced Scorecard implementations across the corporate and non-corporate organizational landscape.

This then begs the question: What is the Balanced Scorecard? What features of its structure and implementation form the conceptual core without which one cannot hope to implement it in its true form and essence? How does one navigate through the various models and variations of the Balanced Scorecard and determine whether or not a modification or extension of the original framework is legitimate or not? To be fair, authors have suffered from similar intellectual conundrums as they looked at other management approaches in the past. Hackman and Wageman (1995) asks similar questions of total quality management (TQM) as they begin to undertake the task of assessing its impact on

[15] Kaplan, Robert S., Presentation at Balanced Scorecard Collaborative Executive Conference, Cambridge: Massachusetts, December 2002.

organizational performance. Yet, one cannot hope to begin this academic exploration without addressing the above concerns and arriving at a reference model that may be used for analysis.

For example, asking whether the Balanced Scorecard is an appropriate performance measurement framework for R&D organizations requires agreeing upon which and whose Balanced Scorecard to use as a reference model. Owing to the importance of setting up an appropriate reference model for the Balanced Scorecard against which we would assess the notions of completeness, performance, effectiveness, and appropriateness, a choice among the multiple (sometimes competing) frameworks is necessary. This is a choice no academic study of this kind would feel comfortable making but one without which none could be conducted at all.[16]

In the light of the above analytical dilemma, and following the tradition of Hackman and Wageman (1995), we make the most obvious and natural choice. We use the version of the Balanced Scorecard promoted by its founders—Drs. Robert Kaplan and his associate David Norton—as the reference for our study. In the subsequent analysis, this version of the Balanced Scorecard would be referred to as Kaplan & Norton's Balanced Scorecard, or the reference Balanced Scorecard, or traditional or original Balanced Scorecard while any variation of it would be referred to as "a" Balanced Scorecard (implementation) or by the nomenclature used by its proponents. As discussed above, this is a necessary choice that we have to make to move ahead with this investigation and while any (and all) of the several competing frameworks is somewhat arbitrary, using any other alternative as a reference would be even more arbitrary than using Kaplan and Norton's Balanced Scorecard. With the issue of a reference framework or benchmark discussed, we are ready to move to define what constitutes a Balanced Scorecard.

3.1— THE BALANCED SCORECARD: THE ORIGINS, MOTIVATIONS & EVOLUTION

While multi-attribute performance measurement frameworks have been in vogue before or around the launch (circa 1992) of the original Balanced Scorecard (e.g. Maisel's Balanced-

[16] The converse of making this choice is to acknowledge that the Balanced Scorecard is so many things to so many different people and that every one of these different frameworks is a legitimate benchmark for us to compare against. This perspective clearly would deprive us of any notion of completeness, effectiveness, and appropriateness as every framework or instance of implementation is unique and thus legitimate, complete, effective, and appropriate in its own right.

Scorecard Model; McNair et al.'s Performance Pyramid; and Adam and Robert's EP2M Model; among others,)[17] the origins of the Balanced Scorecard can be traced back to a 1992 Harvard Business Review article by a Harvard Professor, Dr. Robert S. Kaplan, and a consultant, Dr. David P. Norton, titled ***"The Balanced Scorecard—Measures that Drive Performance"***. Organizational performance, the authors argued, must be measured in a balanced manner that takes into account the various dimensions of performance critical to an organization's success.

Kaplan and Norton (1992) proposed a generic framework that could be applied (as is, or with modifications) to a broad class of business organizations. This generic framework categorized critical dimensions of organizational performance as a set of four organizational performance perspectives, namely, the financial perspective, the customer perspective, the internal process perspective, and the learning and growth perspective. In encouraging organizational leaders and managers to think about organizational performance from multiple "perspectives", the originators of the Balanced Scorecard invoked another important insight. Not only did monitoring financial performance alone misses other dimensions of organizational performance, it also sacrificed long-term vision, strategy, and health of the organization over day-to-day operations and short-term control.

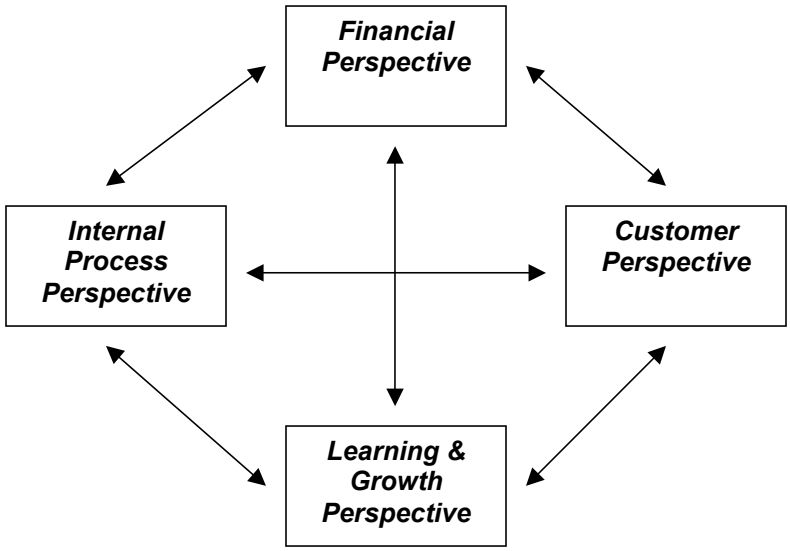


Figure 3.1: Performance Perspectives of a Balanced Scorecard

Standard financial and accounting metrics, it was argued, represent a picture of an organization's past or, at the very best, its present. Citing examples of organizations that failed immediately after being recognized for their financial results, the authors encouraged organizational leaders and managers to measure and monitor the future health of the organization in addition to its current financial performance. Balanced Scorecard provides a

[17] Maisel (1992), McNair et al. (1990) and Adam and Roberts (1993) respectively.

disciplined framework to help them do that by encouraging them to take into consideration key non-financial indicators like customer and brand loyalty and expected future returns from the vitality of R&D pipeline in addition to the traditional financial measures (Olve et al. 1999.)

Although Kaplan and Norton's was only one of the several competing models propagating multi-attribute performance assessment at that time[18], it has become, by many objective and anecdotal standards, the most dominant of this class of performance measurement frameworks[19]. A significant factor in the popularity of the Balanced Scorecard has been the steady intellectual development and evolution of the concept itself. With the idea picking up some currency in the corporate world, the originators of the Balanced Scorecard, and their followers as well as critics, had the opportunity to experiment with and implement the Balanced Scorecard in several hundred organizations around the world. This wealth of experience, once aggregated and synthesized, further led to the evolution of the concept itself.[20] Over the last decade, the Balanced Scorecard has evolved from a system of key

[18] Olve et al. (1999 pp. 20-22) talks about at least three other models that resemble the Balanced Scorecard to varying degrees. Maisel's Balanced-Scorecard Model presented in the Summer of 1992 not only had the same name, but also four performance perspectives—only one of which was different from Kaplan and Norton (1992). Maisel (1992) named the fourth perspective as Human Resources Measures. McNair et al. (1990) presented the Performance Pyramid that incorporated some of the same insights that were behind Kaplan and Norton's Balanced Scorecard. McNair et al.'s Performance Pyramid divides a company into four different levels and provides a structure for two-way communication that is needed to link organization's overall vision and strategy with its operations. The model has separate categories of measures relating to financial/market, customer satisfaction, productivity, and operations etc. Finally, Adam and Roberts (1993) presented the EP2M (Effective Progress and Performance Measurement) model. This model looks at what a company does in four areas, namely, external measures (serving customers and markets), internal measures (improving effectiveness and efficiency), top-down measures (managing strategy and change), and bottom-up measures (empowering ownership and freedom of actions).

[19] Since its introduction, the Balanced Scorecard has become almost a movement in the corporate world—a new management paradigm, so to speak. Kaplan and Norton's 1996 book on the subject has been translated into nineteen languages. Tens of consulting practices exist to help clients implement a Balanced Scorecard. A Balanced Scorecard "Hall of Fame" has also been established by Kaplan and Norton to recognize those organizations that have implemented the Balanced Scorecard in its true spirit and have achieved breakthrough results (Kaplan and Norton, 2001b). The appeal and utility of the idea has been so well received, it is claimed, that the Editors of the Harvard Business School Publishing have hailed it as one of the seventy-five most influential ideas of the twentieth century.

[20] The preface of Kaplan and Norton's 1996 book "The Balanced Scorecard: Translating Strategy into Action"—the first on the subject—encapsulated the essence of this intellectual movement when it stated that "the book, while as comprehensive and complete as we could make it, is still a progress

performance indicators (KPIs) aimed at providing a balanced view of the organization's performance to an organizational communication system (Niven, 2003) and a strategic management system (Kaplan and Norton, 1996.)

A typical implementation of the Balanced Scorecard begins at the top-most level within the organization. Kaplan and Norton (1996) equate the impetus for such an effort with the identification of a "burning platform"—a need-to-change that can no longer be ignored—within the organization. Whatever might be the determinants of the burning platform, it helps create a consensus within the organization about doing things in a fundamentally different way and re-thinking the organizational strategy.

The Balanced Scorecard's four perspectives provide a lens through which the organizational leaders look at the organizational objectives, strategy, processes, and capabilities. In a consultative and participative process, organization's strategic objectives are identified and discussed, the relevant stakeholder segments are identified and specialized strategies for each segment are drawn, organizational processes to implement these strategies are reviewed and (often) put in place afresh, and the various capabilities, skills, equipment, and infrastructure to enable the organization to perform these processes are developed. A chain of causality running from capability to internal process to customer satisfaction to (financial) objectives is thus established throughout the organization.

The participative nature of this strategy-making process ensures that various ideas about potential strategies to achieve organizational objectives would be brought to the table and thoroughly discussed, assumptions underlying actions and initiatives supporting a particular strategy would be questioned and optimized, and an organization-wide awareness of strategy and consensus would be developed that would lead to greater focus and ownership of the organization's strategy. Initiatives, milestones, and performance targets are identified for each sub-segment of the strategy as well as the strategy as a whole. Often performance measurement and data collection systems are put in place to "measure the strategy".

Once the top-level strategy has been developed and agreed upon, a subsequent layer of strategy-making ensues whereby each organizational sub-unit creates its own strategy for achieving its obligations towards the larger organizational strategy. As the process cascades down the organizational hierarchy with multiple levels of Balanced Scorecards being

---

report…we are confident…that innovative companies…will expand the structure and use of the scorecard even further…" (Kaplan and Norton, 1996)

developed and deployed, alignment between the goals of various sub-organizational units and that of the larger organization is achieved. Once this process of strategy-making and scorecards development has taken place throughout the organization, each sub-unit's actual performance against its own performance targets is measured and monitored on a regular basis. Organizations' Balanced Scorecards are often used as a communication tool at management and board meetings thus focusing the discussions on measuring the performance (towards achieving) of organizational strategy. This process also leads to learning across the organization as discrepancies between achieved and targeted performance are detected, their causes are analyzed, and possible remedial actions are suggested.

While the above narrative broadly describes a typical structure and implementation of the Balanced Scorecard, it falls short of incorporating several subtleties and intricacies of the concept that are as essential to the integrity of the framework as these broad details[21]. What is needed is a comprehensive framework that incorporates all necessary ingredients of the Balanced Scorecard and can be used, in its specific or generic form, as a benchmark to compare against. We turn to that task in the next section.

3.2— BENCHMARKING THE BALANCED SCORECARD—AN ANALYTIC FRAMEWORK:

We begin the process of developing a reference model of the Balanced Scorecard by using two constructs, namely, structural features and derived features.

- A ***structural feature*** of the Balanced Scorecard, or for that matter any reference performance measurement system, might be one that is, as the term suggests, built into the structure of the system itself. It is often non-obvious ex-ante, although with the benefit of the hindsight, it may appear to be one of the several logical extensions of the system. At the most basic level, however, it is a characteristic of the way a Balanced Scorecard (or another performance measurement system) is structured. Structural features may be linked with—and are often built on top of—each other. We use the terminology ("SF") to refer to the structural features in the following text.
- A ***derived feature***, on the other hand, is derived from the structure of the scorecard itself. It may result from one or more of the various structural features of the

[21] Kaplan and Norton often describe the Balanced Scorecard framework as "simple, but not simplistic" implying that while the concept seems quite simple, the devil is in the details and implementing it well requires attention to several intricate and subtle ideas and details—a fact that is often overlooked by practitioners and consultants alike.

> Balanced Scorecard. The derived features mostly represent the results and benefits that might automatically arise from putting a performance measurement structure in place. Each derived feature may directly be a result of one or a combination of more than one structural or other derived features. Each structural feature, alone or in combination with other structural features, may give rise to one or more derived features. We use the terminology ("DF") to refer to the derived features in the following text.

An example might help illustrate the difference between a structural feature and a derived feature. The most important structural feature of the Balanced Scorecard is the one used to address the problem of performance multi-dimensionality i.e. it achieves balance across various dimensions of organizational performance by explicitly defining strategies and initiatives along four performance dimensions, namely, financial, customer, internal process, and learning and growth perspectives. Although the four-category scheme appears quite intuitive ex-post, it is not the only logical way to organize a performance measurement system. Indeed, other performance measurement architectures e.g. Six Sigma, TQM, and MBO etc. have been organized differently. The multi-dimensional view of performance, therefore, is a key structural feature of the Balanced Scorecard-type performance measurement system.

What does this multi-dimensional organizing framework of the Balanced Scorecard do for the organization in question? First of all, built into its structure is the fact that it focuses the organization's performance measurement architecture on the achievement of its strategy. Simply put: it "measures the [organization's progress towards the achievement of its] strategy". Putting strategy at the center of the organization's performance measurement architecture, therefore, is a derived feature of the manner in which the multi-dimensional performance architecture of the Balanced Scorecard is set up. It is a derived feature in the sense that it happens as result of a structural feature.

These two constructs, structural and derived features, provide us a way to summarize the various aspects of the Balanced Scorecard in a complete and comprehensive manner. Figure 3.2 presents a snapshot of the various structural features of the Balanced Scorecard. Figure 3.3 presents a tabular representation of the structural and derived features of the Balanced Scorecard. A brief description of each of these is as follows:

3.2.1—SF1: The first key structural feature of the Balanced Scorecard is none other than the notion of balance itself—balance between various performance dimensions (e.g. financial vs. customer vs. internal process vs. employee,) balance between timeframes (e.g. short term vs. long term health,) and balance between types of metrics (e.g. input vs. process vs.

output vs. outcome,) balance between their informational value (e.g. leading vs. lagging indicators,) and balance between perspectives of various stakeholders (e.g. employees vs. customers vs. funding agencies.) The "balance" in the Balanced Scorecard is achieved by identifying four dimensions of performance (the "perspectives")[22]:

- The ***financial perspective*** looks at what the organization needs to do in order to meet the expectations of its shareholders, creditors, security analysts etc. Kaplan and Norton (1996) defined three generic organizational strategies, namely, growth, sustain, and harvest within the financial perspective.[23]
- The ***customer perspective*** looks at what the organization needs to do in order to meet the expectations of its targeted customer segments. It links the key customer-related outcome measures, namely, satisfaction, loyalty, retention, acquisition, and profitability—to its targeted customer and market segments. Kaplan and Norton (1996) introduces[24] the notion of the customer value proposition and identifies three core customer strategies out of which an organization may choose from.
- The ***internal-business-process perspective*** looks at the processes an organization needs to meet its customer and shareholder expectations. The Balanced Scorecard forces organizational managers to specify the complete value chain of internal

[22] A more detailed treatment of the perspectives is done in Kaplan and Norton (1996, 2003) etc. We would very briefly touch upon these to identifying the salient features without going into much detail. The interested reader is advised to look into the above references for more details and examples.

[23] Kaplan and Norton (1996) arrive at three generic themes, namely, revenue growth and mix, cost reduction/ productivity improvement, asset utilization and investment strategy, that might drive the business strategy.

[24] A refined version of the Customer Value Proposition Model was presented earlier by Michael Treacy and Fred Wiersema in The Discipline of Market Leaders (Perseus Books, 1995). It defines customer value in terms of its three components [Value = Product/Service Attributes + Image + Relationship] and drivers that affect each. It consisted of three core customer value strategies, namely, Product Leadership, Customer Intimacy, and Operational Excellence (More on this later.) This model was later adopted by Kaplan and Norton (See Kaplan and Norton, 2001) and assimilated in the "generic" strategy map.

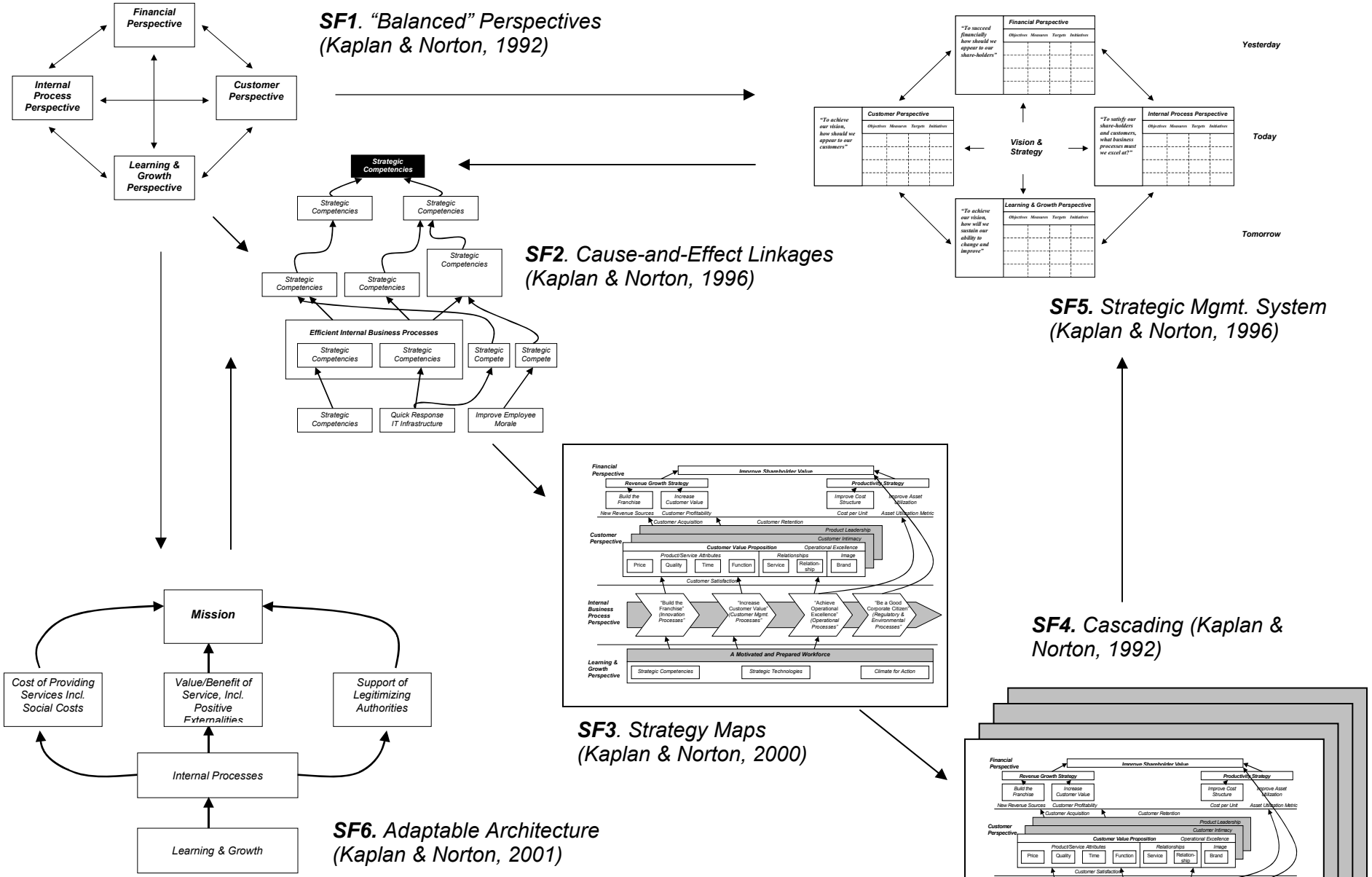


Figure 3.2: The Evolution of The Balanced Scorecard—And Its Key Structural Features

FIGURE 3.3: BENCHMARKING THE BALANCED SCORECARD

STRUCTURAL & DERIVED FEATURES

SF1: Uses perspectives to achieve "balance" between multiple performance dimensions
DF1: Puts strategy at the heart of performance measurement activity
DF2: Emphasizes focused and differential strategies
DF3: Provides an (implicit) model of metrics selection driven by strategy & perspectives
DF4: Provides a framework for undertaking process re-engineering and improvement

SF2: Establish cause-and-Effect linkages between various performance perspectives
DF5: Provides a measurement framework that is meaningful and actionable.
DF6: Provides a means for ascertaining the "correctness" of adopted strategy
DF7: Provides a mechanism for double-loop "strategic learning"

SF3: Provide a mechanism to map an organization's strategy
DF8: Provides a mechanism for simplifying and communicating organization's strategy

SF4: Replicate the Balanced Scorecard throughout the organization by "Cascading"
DF9: Operationalize the organization's strategy into local/individual- level terms
DF10: Provides focus (on strategy) and goal alignment within the organization
DF11: Provides a mechanism for employee involvement and empowerment
DF12: Provides a means for linking incentives and rewards with strategy

SF5: Put Balanced Scorecard at the center of the Strategic Management Framework
DF13: Achieves alignment between and eliminates redundancy among mgmt. functions

SF6: Ensure structural flexibility & adaptability to various organizational forms & circumstances

business processes— starting with the innovation process to operations process to post-sale process—each with its own set of performance indicators and drivers.

- The ***learning and growth perspective*** looks at the quality of the infrastructure—people, systems, procedures (and environment)—necessary to succeed at the objectives set forth in the earlier three performance perspectives. Kaplan and Norton (1996) identify three principal categories of such assets, namely, employee capabilities, information system capabilities, and motivation, empowerment, and alignment—each with its own set of performance indicators and drivers. While this perspective is no less important, it is generally the one least emphasized dimension in organizations (Kaplan and Norton, 1996.)

Although, the primary purpose for using the 4-part categorization of organizational performance is to use metrics from the entire spectrum of organizational activities rather than merely using the financial outcomes, several derived features arise as a result:

- ***DF1: Strategy becomes the cornerstone of the performance measurement activity***. Developing a Balanced Scorecard of measures while thinking through the four interlinked dimensions of performance requires knowledge and awareness of organizational strategy. According to a survey, 85% of the executive teams spend less than an hour a month discussing strategy (Nevin, 2003, p.11) The Balanced Scorecard changes that by putting strategy at the center of the measurement activity. It also facilitates strategy implementation by focusing on "measuring the strategy". (Kaplan and Norton, 2001 p.3)
- ***DF2: Balanced Scorecard emphasizes focused and "differentiated" strategy.*** Most organizations "try to do something for everyone and end up doing nothing for anyone". A strategy-focused organization can either be a product leader (technical excellence,) cost-leader (operational excellence,) or relationship-focused (customer intimacy) but never all of them at the same time (Kaplan and Norton, 2001)[25]. In other words, the Balanced Scorecard forces organizations to have a game-plan to differentiate themselves from their competitors and a focused strategy for doing what they want to do[26].
- ***DF3: Balanced Scorecard suggests an "implicit" model of metrics selection for organizational performance measurement system.*** Although it does not explicitly do so, the Balanced Scorecard implicitly provides a mechanism of metrics selection that often leads to performance indicators that are different from what would have been chosen using traditional methods[27]. The performance measures (generally 10-20 in

[25] Kaplan and Norton (1996) cite the examples of Cray and BMW (Technical Excellence), Dell and Costco (Operational Excellence), and Amazon and (Customer Intimacy) as successful examples of their respective strategies. Organizations generally adopting one of the three customer value propositions must perform satisfactorily on the other two. Trying to do everything for everyone run the risk of failure (akin to NASA's "better-faster-cheaper") and is not a recommended approach.
[26] This idea is so central to the concept of a Balanced Scorecard that one can even questions if a Balanced Scorecard can even be developed without a strategy at all. Organizations would tend to do so, however, they would almost always fall far short of the dramatic performance improvements and strategic alignment that is expected of a true Balanced Scorecard. (Nevin 2003, p.142)
[27] Nevin (2003, pp. 204-6) identifies at least nine other methods, namely, ease of understanding, cause-and-effect linkages, updated frequently, accessibility, average-cautious, resistant to "date"-related measures, quantitative, dysfunction, Balanced Scorecard firmly establishes "link to strategy" as the single most important criteria. If a particular metric doesn't link to the organization's strategy

number)[28] derived from using the Balanced Scorecard are grounded in the organizations' strategic objectives and competitive demands (Kaplan and Norton, 1993) instead of an ad-hoc bottom up process of selecting local performance metrics to measure local activities. Kaplan and Norton (1996) present an alternate view on the number of metrics derived from the Balanced Scorecard[29] in that they view the Balanced Scorecard as the "instrumentation for a single strategy". With that perspective in mind, the actual number of measures becomes irrelevant as they all contribute towards the same single strategy[30].

- ***DF4: It uses strategy as the focal point of an organization-wide process audit and re-engineering.*** Thinking about the internal-business-process perspective requires looking at the organization's business processes with an aim of determining whether or not they are geared towards achieving its strategy as identified in the other three perspectives. This is in stark departure from alternate performance and quality improvement frameworks (e.g. ISO9000, MBNQA) that focus on re-evaluation of existing processes[31]. It forces the organizational leaders and managers to start from a clean slate with a singular focus of equipping an organization with the skill and processes needed to implement its strategy rather than evaluating the existing set of process for possibilities of improvement regardless of their ability to deliver on strategy. This enables the organization to target a breakthrough in its performance rather than settling for a localized performance improvement effort.

---

and doesn't inform the organizational managers anything about it, it has no place in the organization's performance measurement system.

[28] Nevin (2003) provides more detailed estimates of what would be a good number of performance indicators to aim for. With 2-3 strategic objectives under each perspective, and up to 2 (average of 1.5) performance indicators to measure each strategic objective, 15 performance indicators is a good number to aim for.

[29] Kaplan and Norton (1996) invokes the notion of "independent" measures to address the issue: "considering that each of the four perspectives on the Balanced Scorecard can require between four and seven separate measures, businesses often have scorecards with up to 25 measures. Are 25 measures too many? Is it possible for an organization to focus on 25 separate things? The answer to both the questions is NO! If a scorecard is viewed as 25 (or even 10) independent measures, it will be too complicated for an organization to absorb".

[30] Ultimately, the number of performance indicators one should put on an organization's Balanced Scorecard is also a function of what the organizational leaders are trying to achieve with it. A Balanced Scorecard developed for the purpose of merely organizing indicators on an executive's performance dashboard may have greater number of measures than the one used to communicate strategy throughout the organization (Nevin, 2003).

[31] Nevin (2003, p.175) talks about this benefit of the Balanced Scorecard in somewhat detail and compares its approach to organizational improvement with that of other rival frameworks.

3.2.2—SF2: The second key structural feature of the Balanced Scorecard is the cause-and-effect logic that exists between measures and objectives on its various perspectives. A hypothesized or demonstrated cause-and-effect linkages may exist between the financial, customer, internal process, and learning and growth perspectives on the Balanced Scorecard in the sense that one contributes to (or feeds into) the other. To achieve financial objectives, the organization must meet its customer satisfaction targets, which in turn requires it to execute its internal processes well, which in turn requires that relevant skills, equipment, and environment are created within the organization. Kaplan and Norton (1996) recommend going through the above-prescribed sequence while attempting to create a scorecard and paying special attention to thinking about and testing the chain of causality of the hypothesized organizational strategy. This adds value to the overall exercise in several ways:

- ***DF5: It creates a measurement framework that is meaningful, actionable, and adaptable.*** This is done, primarily, by picking a set of performance metrics relating with different stages of strategy execution, namely, input, process, outputs, and outcomes that are causally linked to each other. Another way to characterize this derived feature of the Balanced Scorecard is that it aims to "measure the strategy". Monitoring a mix of leading and lagging indicators makes the exercise of measurement much more real, meaningful, and credible to employees who may not see the ultimate effects of the strategy until several months or years but can use the leading indicators as measures of its ultimate success. Such a measurement framework also makes the exercise of measurement more actionable allowing organizational leaders, managers, and employees to compare observed interim effects (on leading indicators) with expected interim effects (hypothesized effect-sizes of causal linkages) as well as examine the hypothesized relationships between measures across the four perspectives, and take corrective actions or fine-tune the strategy.
- ***DF6: It provides the ability to check the "correctness" of the adopted strategy.*** A strategy, by its very nature, is a plan or a "hypothesis" (Porter, 1996) that organizational leaders and managers believe would take it from point-A to point-B. Implementing a previously untested strategy often takes an organization through uncharted territory. Whether or not a strategy works and actually takes the organization from point-A to point-B not only depends on the quality of its implementation but also on whether or not the hypothesized cause-and-effect linkages were actually correct. Balanced Scorecard makes the linked hypotheses of the strategy explicit and testable. (Nevin, 2003 p. 169, and Kaplan and Norton, 2001

p. 76.) The cause-and-effect feature of the Balanced Scorecard allows organizational leaders and managers to test these hypothesized relationships[32].

- ***DF7: It provides an opportunity to engage in strategic double-loop learning.*** Kaplan and Norton (1996, 2001)—drawing upon the work of Chris Argyris--describe this phenomenon as "double-loop learning" (Argyris, 1991). Double-loop learning refers to the learning that occurs when organizational leaders and managers question their assumptions and reflect on whether the theory under which they were operating (i.e. the hypothesized cause-and-effect linkages) is still consistent with current evidence, observations, and experience (Kaplan and Norton, 1996 p.251)—an option that only becomes available if the strategy is meticulously measured. Balanced Scorecard does that by defining a well-defined and articulated strategy and linking it with the performance measurement system through an implicit underlying performance model. Several companies studied by Kaplan and Norton have taken this learning opportunity even a step further by using formal statistical techniques to supplement their intuitive understanding of the strategy[33]. Based on this feedback, many of the organizations are able to fine-tune their strategies[34].

3.2.3—SF3: The third key structural feature of the Balanced Scorecard is a device called "strategy mapping" that allows the scorecard developers to "tell a story" about the organization's strategy. The strategy maps are graphical devices that represent the strategy of an organization in a concise, graphical, and user-friendly manner. Strategy mapping is a recent addition to the Balanced Scorecard literature—introduced primarily through customer innovation[35]. Kaplan and Norton (2001) suggests using "generic" strategy maps

---

[32] Kaplan and Norton (2001) illustrates this feature of the Balanced Scorecard with the help of several examples. For example, they present a case of a chain of convenience stores that discovered—after the implementation of the Balanced Scorecard—that the hypothesized causal logic in its strategy was faulty. By making explicit the causal connections (or the lack of it), the Balanced Scorecard allowed this particular organization to re-think the strategy and make corrections midcourse without losing too much time and money pursuing a flawed strategy.

[33] When sufficient data are available for statistical analysis, managers have a powerful tool to not only confirm (or reject) the validity of their hypothesized relationships, but also to begin estimating quantitatively the magnitude and time-lags of linkages across measures (Kaplan and Norton, 2001 p.311).

[34] Kaplan and Norton (2001, p. 316) describe this in the following manner: "Organizations find that up to 25% of the measures change every year, reflecting the learning that is taking place. The original strategies or the previous measures were not wrong. The evolution was caused by the organization reaching another level of sophistication around its strategy."

35 This conjecture can be verified from the fact that Kaplan and Norton's 1996 Book "Balanced Scorecard: Translating Strategy into Action" does not talk about "Strategy Mapping" in any

(please see Figure 3.4) as a device to initiate brainstorming and organizational dialogue on strategy itself. A "generic" strategy map also formalizes the notion of generic strategic themes for each perspective of the Balanced Scorecard. A well-designed Balanced Scorecard should "tell the story" of the organization's strategy" (Kaplan and Norton, 2001). Ideally speaking, a well-done strategy map would allow an individual—an employee or other stakeholder—to deduce and decipher an organization's strategy through a little bit of thought and reflection. Best-told stories allow an organization's employees to see what the organization is trying to accomplish and how their work fits into the overall picture (Nevin, 2003). There are several other features that are derived from strategy-mapping:

- ***DF8: It helps simplify and communicate organization's strategy to its employees and other stakeholders.*** Kaplan and Norton (2001) identify a lack of understanding of strategy within organizations as a major barrier to implementation of strategy and achievement of breakthrough performance improvement. Only 5% of employees within organizations understand strategy (Nevin, 2003). However, understanding the strategy is critical to executing it well[36]. The Balanced Scorecard, by demystifying the organization's strategy through the use of strategy maps, attempts to rectify this situation. By communicating the organization's strategy at both employee and senior manager levels[37], it moves strategy from the Boardroom to the Backroom and makes it more accessible to an average back office or frontline employee[38].

3.2.4—SF4: The fourth key structural feature of the Balanced Scorecard is the process of cascading that refers to replication of the Balanced Scorecard throughout the

---

substantive way. While it does contain figures with boxes and arrows running through the four Balanced Scorecard perspectives, the index does not contain the word "Strategy Map" itself. Contrast this with the title of the HBR article by Kaplan and Norton titled "Having Trouble with Your Strategy: Then Map It" (Sept-Oct. 2000) and their book "Strategy-Focused Organization" (2001) that dedicates an entire chapter to the subject and tens of references to it in the index.

[36] Kaplan and Norton (2001, p.65) describe the importance of understanding of strategy among an organization's employees in the following manner: "In the era of knowledge workers, strategy must be executed at all levels of the organization. People must change their behaviors and adopt new values. The key to this transformation is putting strategy at the center of the management process. Strategy cannot be executed if it cannot be understood, however, it cannot be understood if it cannot be described."

37 In the words of Kaplan and Norton (2001, p. 104): "We do not claim to have made a science of strategy. The formulation of strategy is an art, and it will always remain so. The description of strategy, however, should not be an art."

[38] At one organization they helped develop a Balanced Scorecard for, Kaplan and Norton (2001 p.218) noted that only 20% of the employees were aware of the organization's strategy before the implementation began—a number that stood at around 80% five years later once the Balanced Scorecard was rolled out throughout the organization.

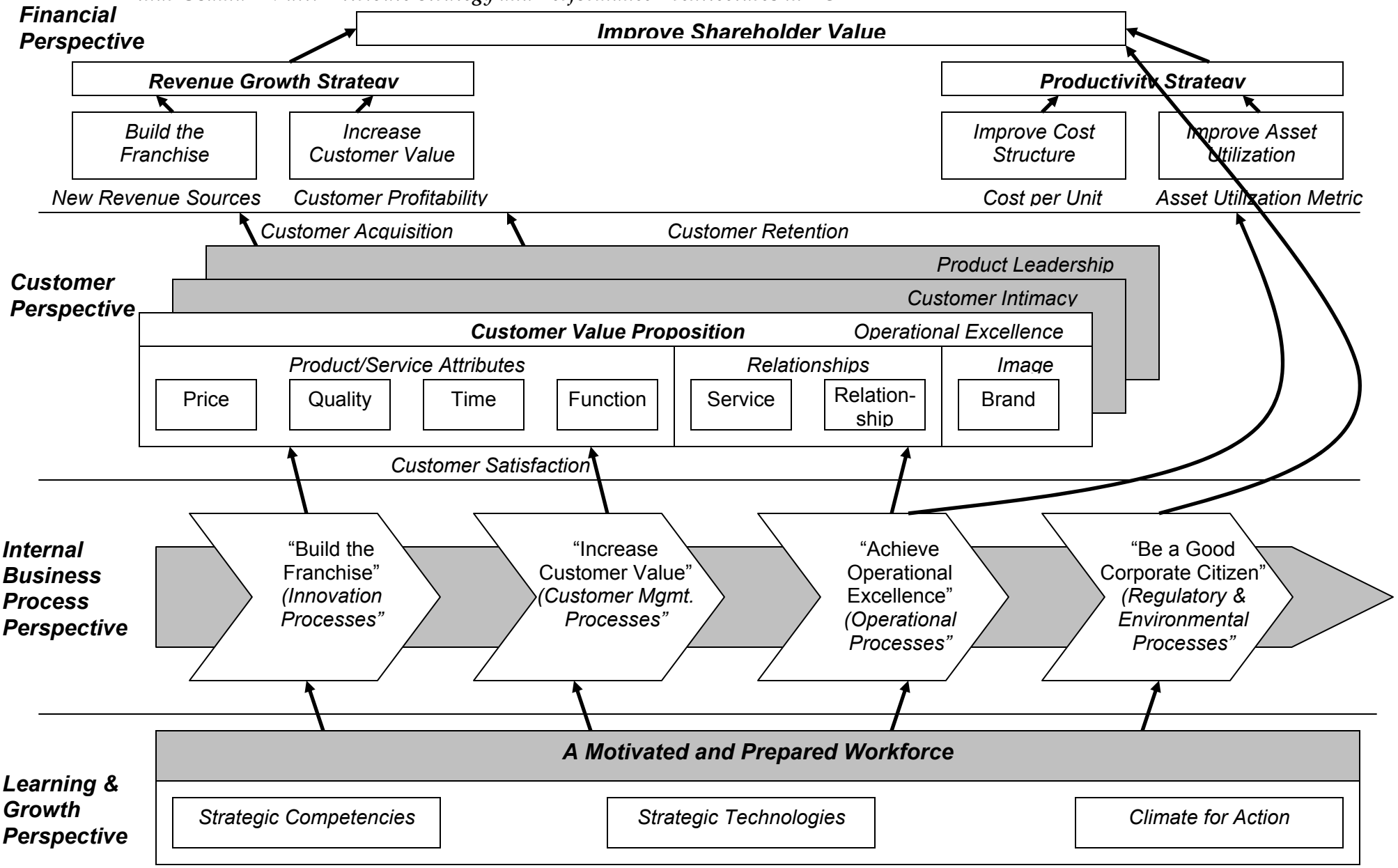


Figure 3.4: A Generic Balanced Scorecard Strategy Map (Source: Kaplan & Norton (2001, p. 96)

organizational hierarchy. The key benefit of cascading is aligning the goals, objectives, and strategies of various sub-organizational units with that of the overall organization. Even when Balanced Scorecards have been developed at the top levels of the organization, significant "learning gaps" exist within the organization not only in clarifying how individuals' jobs fit into the overall organizational picture but, more importantly, in clarifying what the overall organizational picture really is? This is achieved through the communication, transparency, and, most importantly, the process of cascading. Kaplan and Norton (2001 p.45) suggest that lower level scorecards be developed in the light of local conditions. Measures at the individual business levels do not necessarily have to add up to a higher-level measure. However, it is necessary that local measures be chosen in a manner that they influence the measures on the higher-level scorecard[39]. The importance of cascading in driving the organizational strategy and the Balanced Scorecard home to the vast majority of an organization's employees cannot be over emphasized[40]. Additionally:

- ***DF9: It operationalizes the organization's strategy to its employees and makes strategy "everyone's job".*** In layman terminology, operationalization of the strategy refers to the translation of the organization's overall strategy—often the "big picture" of where an organization wants to go—into its constituent sub-organizational and even individual components. Many a times, and to varying degrees, this process happens in a decentralized manner in the minds of the employees as they see the linkage between where the organization is heading and how their work relates to it. Cascading helps formalize this process by forcing an organization to successively implement the Balanced Scorecard throughout the organizational hierarchy down to, at times, the level of an individual[41]. This makes strategy accessible and meaningful to each and every individual

[39] Kaplan and Norton describe this in the following way: "In general, it is not critical or even desirable for all measures on the lower-level scorecards to add up to measures on the high-level scorecard. We have found it preferable for the lower-level unit to have discretion to customize the measures to their own circumstance, rather than receive a drill down component from a higher-level scorecard." (Kaplan and Norton, 2001 p.174)

[40] Nevin (2003) stresses upon the importance of the cascading in the following way: "Of those items within your control, cascading may be the single most important ingredient of successful Balanced Scorecard implementation. You cannot control how much executive sponsorship you receive or predict any crisis that might derail your efforts. You can, however, make the decision to drive the of the Balanced Scorecard to all levels of your organization."

[41] When individual can construct their own Balanced Scorecards, then two important things happen. Firstly, they clearly understand how they contribute towards the achievement of organization's strategic objectives and hold themselves personally accountable for doing their part, and secondly, the organization has developed the clearest mechanism for aligning individual objectives to business unit and corporate objectives. (Kaplan and Norton, 2001 p.244)

within the organization. Once employees can relate with strategy, they can begin to implement it.

- ***DF10: Balanced Scorecard provides laser-like[42] focus on strategy and goal alignment within the organization.*** In many ways, the focus and alignment is key to the performance breakthroughs that have been reported by organizations implementing the Balanced Scorecard (Kaplan and Norton, 2001). Yet, this is something that is often not done as well as it should be[43]. A rich case study literature—although anecdotal—provides some evidence on the ability of the Balanced Scorecard to set into motion a decentralized process to achieve this focus and alignment. In several cases, employees realized how their work fits in with the work of other people and how they could—above and beyond what they were already "required" to do—think of clever ways to create more value to enable their organizations to achieve its stated goals and strategic objectives.
- ***DF11: Balanced Scorecard encourages employee involvement and empowerment vis-à-vis strategy formulation and implementation.*** With knowledge of strategy comes involvement and empowerment. Only when employees within the organization know what the organization's strategy is, and they understand how they contribute to it, can they begin to feel involved and a part of the strategy-making process[44]. Kaplan and Norton (2001, p.48) suggest using the Balanced Scorecard to communicate strategic objectives to the employees, but not to command them what to do. In a sense, the Balanced Scorecard thus becomes a mix of top-down and bottom-up strategy formulation and implementation mechanism. The result often is enthusiasm, initiative, and ownership—and an organization committed to implementing a strategy.

---

[42] Kaplan and Norton (2001, p.8) describes this focus and alignment with the help of the example of diffused light and the focus of the laser. "Think of diffused light produced in a well-lit room by thousands of watts of incandescent and fluorescent lamps. Compare that warm light with the brilliant beam of light that comes from the tiny battery in a handheld laser pointer….the laser operates non-linearly; it leverages its limited power source to produce an incredibly bright and focused beam of light. Similarly…a well-crafted and well-understood strategy can, through alignment and coherence of the organization's limited resources, produce a non-linear performance breakthrough."

[43] Consider the following data from Kaplan and Norton (2001, p. 234):

- 51% of senior managers in US, 31% in UK have their personal goals linked with strategy
- 21% of middle managers in US, 10% in UK have their personal goals linked with strategy
- 7% of line employees in the US, 3% in UK have their personal goals linked with strategy

[44] This is sometimes referred to as akin to capturing the hearts and minds of the organization's employees. When employees get involved in strategy-making, they also own it. With ownership of strategy comes the aligning of hearts. Once they own the strategy, they want to begin thinking about how to implement it as well. This releases the combined creativity energy of the organization's employees in support of the organization's strategy—the alignment of minds, so to speak.

- ***DF12: Cascading provides a mechanism of tying individual incentives and rewards with their performance.*** Linking incentives and rewards with quantitative measures of performance is often viewed as quite controversial[45] (Austin, 1996,) yet despite the perceived difficulties in doing so, it is often practiced in one form or the other. Regardless of what side one is on this debate, the only generalizable finding from several company experiences in linking compensation and reward to Balanced Scorecards is that they do it (Kaplan and Norton, 2001 p.265.) Cascading the Balanced Scorecard down to each and every employee (or workgroup)—with each personal scorecard being linked to the successive higher-level scorecards—provides a means to focus individual attention on strategy[46] and align personal goals with organizational objectives, thus providing a clear way to link individual incentives and rewards with their performance.

3.2.5—SF5: The fifth key structural feature of the Balanced Scorecard is its ability to establish a strong connection between performance measurement and components of the larger strategic management system. While originally not conceived to be a strategic management system, the Balanced Scorecard, through sheer customer-driven innovation and feedback, has ended up becoming one[47]. The most important feature of the Balanced Scorecard that drives this transformation is its ability to put organizational strategy at the center of the debate on what an organization does, and how should its performance be measured. Kaplan and Norton (1996b) describe a system of strategic management (or a system for managing strategy) as having four components, namely, an ability to a) translating the vision and strategy into actionable and operational objectives and measures, b) communicating the vision and strategy to all employees within the organization and linking rewards to performance measures, c) aligning business

[45] The proponents of "pay-for-performance" highlight the need for extrinsic factors in shaping individuals' behavior. The critics argue that linking real incentives and rewards to "numbers"—no matter balanced they are—leaves much to be desired. It creates the possibilities of dysfunctional behaviors like fudging the numbers and/or misguided focus on what being measured rather than what is necessary.

[46] An interesting perspective on the subject comes from a senior executive at Mobil North America, who when colleagues chided him for how much time his people spent time studying the scorecard results each month, responded: *"I think just the opposite. Is fabulous that people spend that much time. For one each month, all employees are taking out the scorecard and looking at the most important things in their business and whether we were winning or losing against targets. They were doing this to see how much money they were going to get. We would not have gotten the same focus on the scorecard, and individual business objectives, if we hadn't linked it to pay"* (Kaplan and Norton, 2001 p.255).

[47] Kaplan and Norton (1996b) describes Balanced Scorecard's transformation from a key performance indicator (KPI) system to a strategic management system in the following words: "Several years ago…we introduced a concept we called the Balanced Scorecard…Recently, we have seen some companies move beyond our early vision for the scorecard to discover its value as the cornerstone of a new strategic management system." (p. 75)

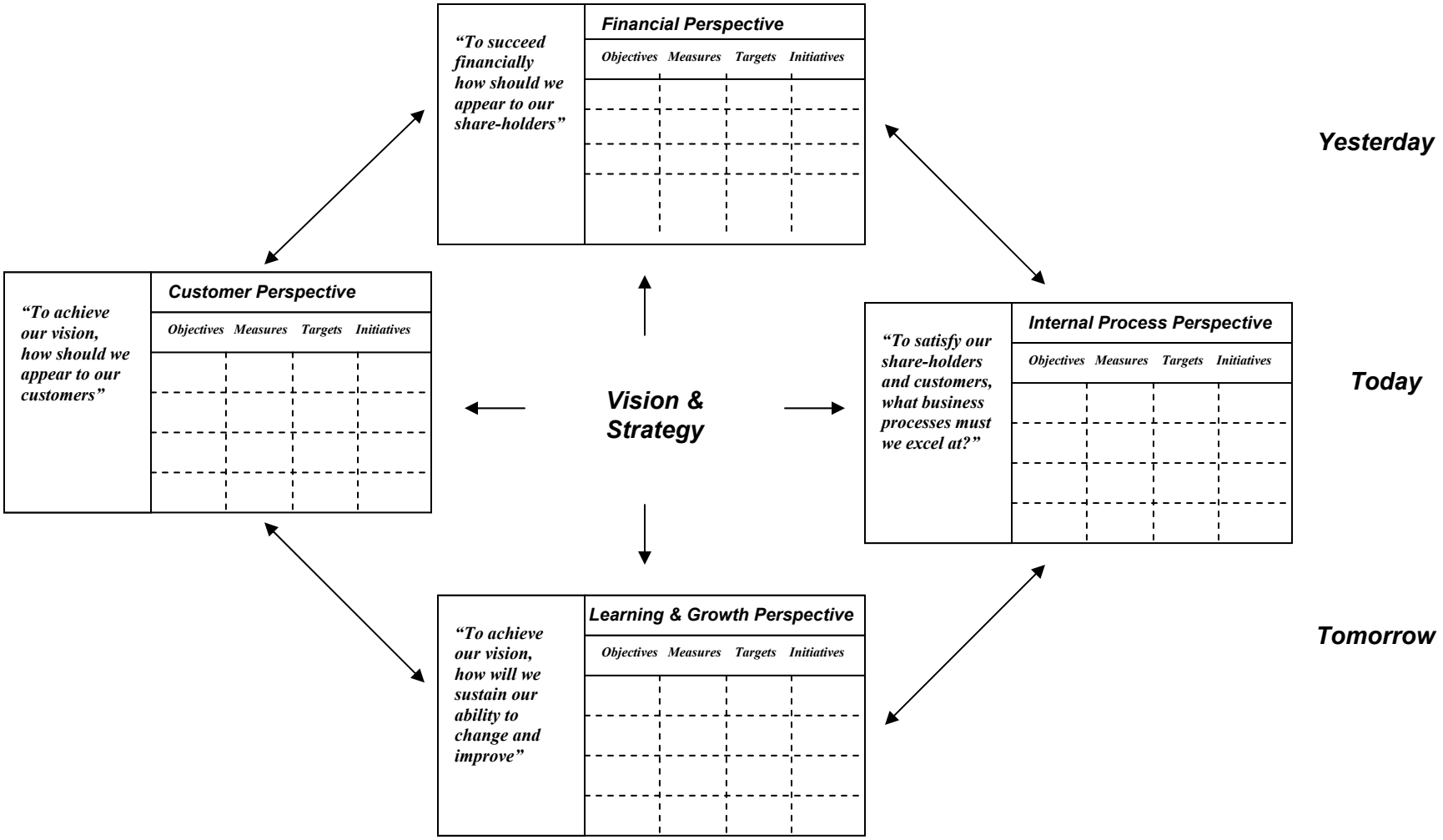


Figure3.5: Graphical Representation of the Balanced Scorecard. (Source: "Using the Balanced Scorecard as a Strategic Management System" by Robert S. Kaplan and David P. Norton, January-February 1996, Harvard Business Review, p.76)

planning functions with vision and strategy (e.g. setting targets, identifying initiatives, allocating resources etc.), and d) facilitating strategic feedback and learning.

By using the Balanced Scorecard as a unifying framework with which other components of a strategic management system seamlessly connect, the possibility of a major organization dysfunction, namely, a disconnect between performance measurement and the larger strategic management system, is eliminated. For example, it is possible that the organizational budgeting process—for the reason of ensuring stability of funding—may not reflect changing organizational priorities and strategy. This is often the case when budgetary process is carried out in isolation with the strategy-formulation process. By requiring that initiatives, budgets, and milestones be identified to execute each element of the Balanced Scorecard strategy, it eliminates this possibility (please see Figure 3.5)[48]. Similarly, the Balanced Scorecard ties together other elements of a strategic management system e.g. organizational communication, learning, compensation etc.

- ***DF13: Balanced Scorecard achieves alignment and eliminates redundancy (and waste) between managerial processes and systems.*** This is a powerful statement to make about a management framework of any kind. Balanced Scorecard's ability to provide insights and guidance on, and link together all major corporate managerial functions into a seamless whole, on the one hand, and its integral connection with corporate strategy, on the other hand, can be a great source of alignment and coherence in an organization. A central organizing framework for managerial functions—like the Balanced Scorecard—would result in lesser confusion in the organization as employees and managers would know exactly the underlying management logic of the organization (Kaplan and Norton, 1996 p.283.) It would also help reduce redundancies and waste from managerial processes and ensure that they do not result in actions that are counterproductive to each other.

3.2.6—SF6: The sixth key structural feature of the Balanced Scorecard is its flexibility and adaptability to varying organization types, forms, and situations. Flexibility and adaptability, although not germane to the structure of the Balanced Scorecard itself, is central to its success as a management framework. Adaptability is possibly the Balanced Scorecard's greatest attribute (Nevin, 2003 p. 182). This flexibility manifests itself in many

[48] Kaplan and Norton (2001) suggests the use of so-called dynamic budgeting—a technique that represents a mix of operational and strategic budgeting. They suggest using a variant of activity-based costing approach to allocate resources to sustain operational activities with the Balanced Scorecard-linked strategic initiatives receiving a full budgetary review. Navin (2003) provides templates for isolating the operational and strategic component of each Balanced Scorecard initiative and critically looking at it.

ways. First is the flexibility in the basic structure of the scorecard. While the originators of the Balanced Scorecard proposed looking at four "perspectives", they also kept the structure flexible enough to incorporate re-naming, re-ordering, and addition or subtraction of the perspectives prescribed by them.[49] Many have re-named the tool itself to make it appear more aligned to their organization's culture.

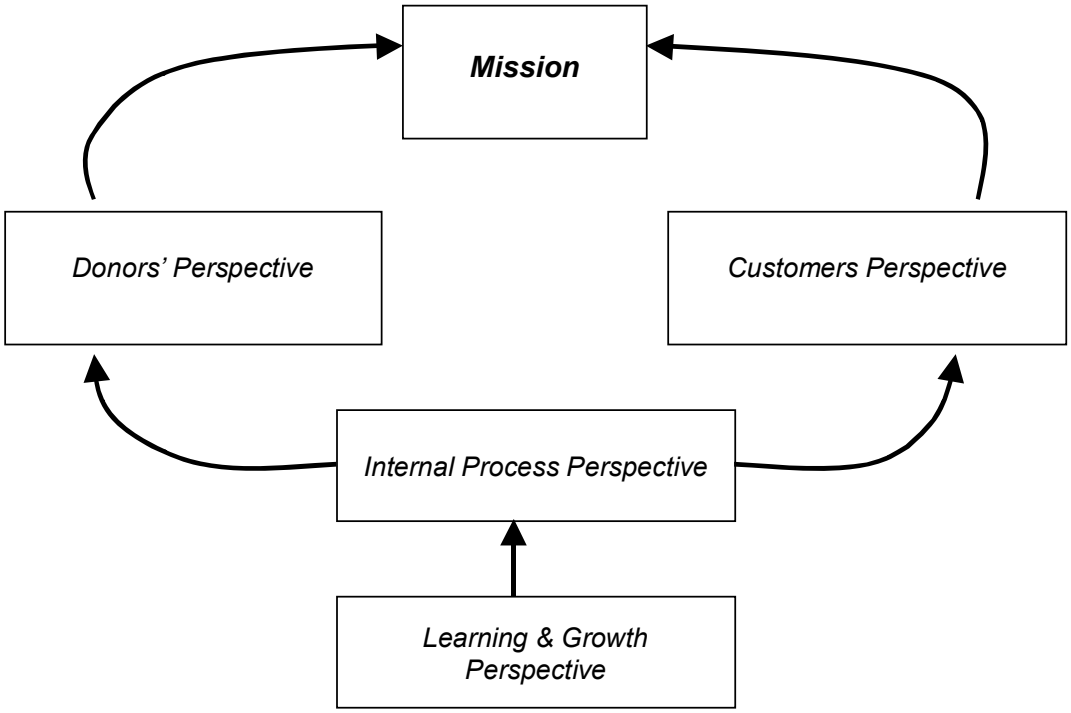


Figure 3.6: A Modified Balanced Scorecard Template for Non-Profit Organizations (Adapted from Kaplan and Norton, 2001 p. 135)

While a number of public sector agencies use the same four perspectives on their balanced scorecard (Nevin, 2003 p. 182), many have experimented with variations. Mission-driven non-profits and public sector organizations have experimented with using "Mission", in place of financial perspective, as a dimension that represents the bottom-line for their organizations. Many non-profits and public sector organizations have also used a "split" customer perspective to represent their multiple customers— direct (e.g. beneficiaries of their services) and indirect (e.g. funding agencies, contributors, the Congress) Kaplan and Norton (2001, p.136), for example, present a modified version of the Balanced Scorecard that rearranges the perspectives to cater to the different requirements of non-profits (please see Figure 3.6) and public sector agencies (please see Figure 3.7).

[49] The Dallas Family Access Network—a non-profit healthcare organization—for example, adds a "healthcare perspective", a "social services perspective", and a "consumer" perspective to the usual operational and financial perspectives.

The above structural and derived features provide a comprehensive view of the structure and functioning of the Balanced Scorecard. Many of these features are a result of practitioner-driven innovation. Hundred of organizations that have implemented the Balanced Scorecard have found one (or more) of its features more useful than others and have experimented with using it in ways that were totally unanticipated by its creators in 1992 (Kaplan & Norton, 2001.) In doing so, these organizations have played a part in putting together the missing pieces of this measurement—and later, strategic management--framework. Collectively, each of the above feature contributes to the completeness and effectiveness of the overall framework. For the purpose of subsequent analysis, therefore, we would use this set of features as a benchmark of an effective Balanced Scorecard.

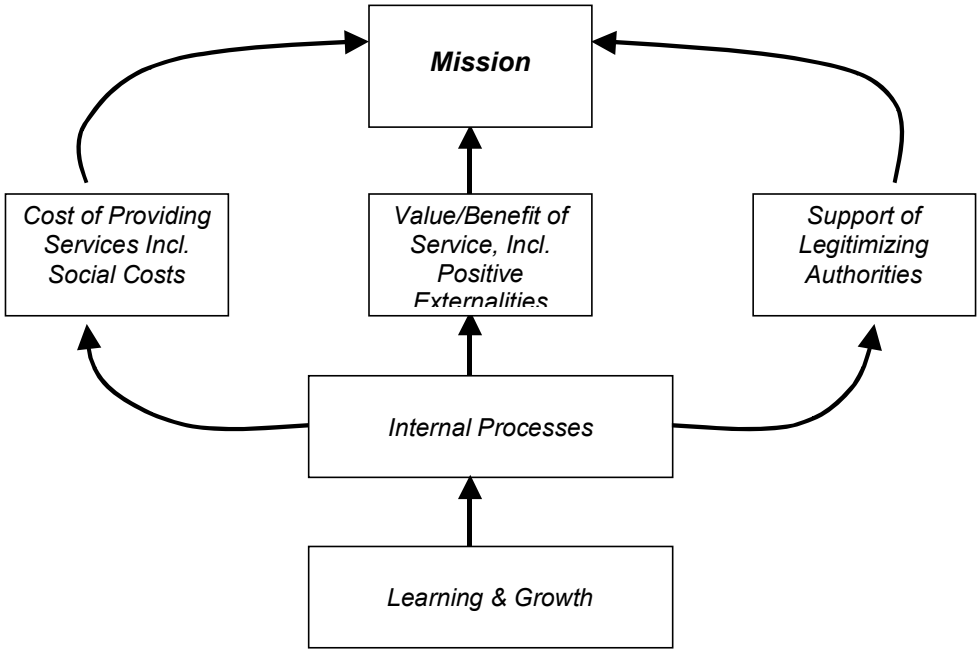


Figure 3.7: : A Modified Balanced Scorecard Template for Public-Sector Agencies (Source: Kaplan and Norton, 2001 p. 136)

Does the Balanced Scorecard tell us something about what constitutes an effective performance measurement and management system? A closer reflection reveals that the structural and derived features of the Balanced Scorecard can be generalized into a set of ideas or principles for effective performance measurement and management systems. In essence, this gives us a generic analytic framework to not only assess the extent and quality of implementation of a Balanced Scorecard within an organization but also as a benchmark against which other performance measurement and management frameworks may be

compared. Before we do that, however, we must translate the six structural features and their derived benefits into generic components, processes, and outcomes. Figure 3.8 presents a set of generic structural and derived features that are a direct translation of the structural and derived features of the Balanced Scorecard.

For example, SF1 relates to the four perspectives that provide the important ingredient of "balance" to the Balanced Scorecard. While the four perspectives provide "a" mechanism to achieve balance between various performance dimensions, it might not be "the" (only) mechanism to do that. Technology Value Pyramid, for example, provides another way of achieving this balance between various dimensions of performance, as does the Performance Prism (Tipping et al., 1995, Neely et al., 2002 respectively.) Some performance measurement frameworks might do this better than others—and some might not do it well at all. Regardless of how other performance measurement frameworks do it, the notion that organizational performance is a multi-dimensional concept and that achieving a balance between various dimensions of performance is desirable and hence makes an important feature of a "generic" performance measurement system.

This insight provides us with a powerful framework for evaluating the "completeness" and "effectiveness" of a performance measurement system by generalizing from the specific (i.e. established literature on experiences of organizations that have successfully used the Balanced Scorecard) to the generic (i.e. features and characteristics of an effective performance measurement system). We may use this framework in at least a couple of ways:

a) First, the structural and derived features maybe used as a checklist against which the state of a Balanced Scorecard implementation in a specific organization may be compared. This checklist maybe used to assess whether a particular organization has indeed implemented the Balanced Scorecard and whether or not it has done so in its entirety (i.e. incorporating the structural features) and in the true spirit, as prescribed by its founders (i.e. realized the derived features).

Secondly, the corresponding generalized structural and derived features (GSF and GDF) may serve as benchmarks against with performance measurement and management systems of other types may be compared. In this format, we may be able to assess whether or not a particular performance measurement and management system is providing all, or a subset of, features that may have been

FIGURE 3.8: A GENERIC PERFORMANCE MEASUREMENT AND MANAGEMENT FRAMEWORK

THE BALANCED SCORECARD	A GENERIC MODEL
SF1: Uses perspectives to achieve "balance" between multiple performance dimensions *DF1: Puts strategy at the heart of performance measurement activity* *DF2: Emphasizes focused and differential strategies* *DF3: Provides an (implicit) model of metrics selection driven by strategy & perspectives* *DF4: Provides a framework for undertaking process re-engineering and improvement*	**GSF1: Adopts a "balanced" view of the organization and its performance** *GDF1: Links what is being measured with what needs to be achieved ("strategy")* *GDF2: Emphasizes "differential" view of strategy* *GDF3: Provides an explicit or implicit model of metrics selection* *GDF4: Provides a "global" (rather than local) model of organizational improvement*
SF2: Establish cause-and-Effect linkages between various performance perspectives *DF5: Provides a measurement framework that is meaningful and actionable.* *DF6: Provides a means for ascertaining the "correctness" of adopted strategy* *DF7: Provides a mechanism for double-loop "strategic learning"*	**GSF2: Explicitly links, through "evidence-based" reasoning, all performance measurement activities together into a Single whole and with the strategy** *GDF5: Provides meaningful and actionable insights to managers* *GDF6: Provides a means to ascertain "correctness" of strategy* *GDF7: Provides a means of "strategic learning"*
SF3: Provide a mechanism to map an organization's strategy *DF8: Provides a mechanism for simplifying and communicating organization's strategy*	**GSDF3: Clearly establishes and communicates the measurement system's link with strategy** *GF8: Provides a mechanism for simplifying and communicating organization's strategy*
SF4: Replicate the Balanced Scorecard throughout the organization by "Cascading" *DF9: Operationalize the organization's strategy into local/individual- level terms* *DF10: Provides focus (on strategy) and goal alignment within the organization* *DF11: Provides a mechanism for employee involvement and empowerment* *DF12: Provides a means for linking incentives and rewards with strategy*	**GSF4: Make strategy-implementation and strategy-measurement an organization-wide, not a "localized" activity** *GDF9: Operationalize the organization's strategy into local/individual- level terms* *GDF10: Provides focus (on strategy) and goal alignment within the organization* *GDF11: Provides a mechanism for employee involvement and empowerment* *GDF12: Provides a means for linking incentives and rewards with strategy*
SF5: Put Balanced Scorecard at the center of the Strategic Management Framework *DF13: Achieves alignment between and eliminates redundancy among mgmt. functions*	**GSF5: Integrate performance measurement within the Strategic Management System** *GDF13: Align various mgmt. functions*
SF6: Ensure structural flexibility & adaptability to various organizational forms & circumstances	**GSF6: Ensure flexibility & adaptability for various organizational forms/missions**

available had a Balanced Scorecard been implemented. This allows us to assess the completeness and the effectiveness of a particular performance measurement technique using the Balanced Scorecard as a benchmark. We would return to this analytic framework, both in its Balanced Scorecard-specific and generic formats, in the subsequent analysis.

3.3—CHALLENGES TO AND WEAKNESSES OF THE BALANCED SCORECARD

While the discussion so far, has focused on the explaining the attributes of the Balanced Scorecard—using the writings and preaching of its founders (i.e. Dr. Robert Kaplan & David Norton's) as a reference model—we would now turn to presenting some of its weaknesses described by the framework's critics. The following discussion would be divided into two parts. The first part would address some specific challenges and criticism of the Balanced Scorecard concept while the second part would address more generic ones that are not germane to Balanced Scorecard alone but may be generalized to other such frameworks and approaches as well.

3.3.1—Specific Challenges to the Balanced Scorecard

Gupta (2004) presents one such critique on the Balanced Scorecard. It uses a mixture of Six Sigma methodology and the Balanced Scorecard for establishing an overall performance level for an organization. While acknowledging that the Balanced Scorecard "reveals a much broader view of what is happening in an organization than traditional financial measures do", it enumerates several of its shortcomings. The most important of these is the measurement burden associated with implementing a Balanced Scorecard all through the organization. "By the time the Balanced Scorecard gets to work groups", asserts the author, "the strategy has become unrelated to employees and too much effort is required to maintain the system". Secondly, the author points out that [Balanced] scorecard-type systems are ineffective in relating to the employees who do the work in the sense that they are "strategic in nature and do not flow down to process measures". (Gupta, 2004, p.15).

It presents its own performance measurement system (i.e. Six Sigma Business Scorecard) as alternative to the Balanced Scorecard that replaces:

- the latter's emphasis on strategic intent with emphasis on maximizing profitability,
- the latter's focus on growth with the focus on profitability,
- the latter's use of targets (i.e. levels) with the use of rates of improvements,
- the latter's aspiration for balanced set of measures with the aspiration for a set of measures that emphasize impact (Gupta, 2004, p.88.)

This is not a place to delve into the argument of whether or not Gupta (2004) presents a better or worse alternative to the Balanced Scorecard. However, the two critiques presented

above merit a review. We find some truth to the suggestion that implementing a Balanced Scorecard across the various levels of an organization can be a resource intensive task. Olve et al. (1999) puts a typical timeframe for rolling out an organization's first (top-level) scorecard at about 3-6 months. Depending on the size of the organization, it might take 2-5 years to develop and implement Balanced Scorecards across the entire organizational hierarchy. There is definitely ample room here for the organization and its employees to lose their sense of direction and track of why they are going with the motions of implementing a Balanced Scorecard.

Experts, therefore, have recommended the use of detailed communication, both upfront and throughout the rollout of the scorecard, as a means to keeping the effort on track (Kaplan and Norton, 2001, Olve et al., 1999, Nevin 2001 etc.) and keeping people focused on the effort. It is also important to aim for and emphasize upon some early gains (the proverbial "low hanging fruit") and to create positive energy and expectations about the exercise as the Balanced Scorecard is rolled out throughout the hierarchy.

Another way to look at this issue, however, is to view a Balanced Scorecard implementation as the implementation of organizational strategy itself. It is probably not feasible or realistic for people to expect an organization's new strategy to deliver within a month, or three, or even six months time. The Balanced Scorecard embodies an organization's strategy and its effect on the performance of an organization is the manifestation of the success or failure of that strategy. What is needed then, to safeguard against the prospect of the exercise losing steam and relevance, is to quickly put in place the Balanced Scorecard's performance measurement architecture while at the same time conveying the right expectations about the timeframe for any improvements in performance to become available and visible.

That Gupta's (2004) second critique on the Balanced Scorecard, i.e. the scorecards are too operational and not strategic enough is not valid, should be evident from the description of the Balanced Scorecard in section 3.2. Specifically, it criticizes the fact that the Balanced Scorecard does not pay enough emphasis on process measures—a fact belied by the presence of internal business process as one of the four key performance dimensions. In fact, "operationalizing the strategy" is a key feature of the Balanced Scorecard and is appropriately recognized in our reference or benchmark model as one of the several derived features (DF9). Gupta's critique, therefore, can only be representative of ill-conceived and implemented efforts but fails to make a dent in the theoretical core of the Balanced Scorecard methodology.

Marshall W. Meyer (2002) presents a more philosophical critique that hits at the core concept of the Balanced Scorecard when it asserts that organizational performance lacks construct validity. Meyer (2002) asserts that organizational performance measures are generally uncorrelated, and that we normally confuse performance measures with performance itself. It goes further to assert that all performance measures are imperfect and it is very difficult to find the "right" measures in an environment as dynamic as an organization's. "I am skeptical of basing strategy on performance measures", he asserts, because "I worry about the unintended consequences".

The Balanced Scorecard, according to this critique, tries to build an entire strategic management system on a bunch of measures without addressing the underlying methodological issues e.g. defining performance, measuring it, finding the right measures. "The [technical] requirements of an ideal performance measurement system are very stringent, far more stringent than the requirements of the Balanced Scorecard", Meyer (2002, p. 7) concludes[50]. Meyer (2002) presents an alternative to the Balanced Scorecard—an approach called Activity-based Profitability Analysis (APBA)—that replaces Balanced Scorecard's arbitrariness[51] by one that picks performance measures using a firm quantifiable construct, namely, sales profitability.

Meyer's (2002) critique on the Balanced Scorecard is somewhat, though not entirely, valid. For example the assertion that the Balanced Scorecard does not provide a bottom-line for the organization, and Meyer's ABPA does, is found to be not quite true on a moment's reflection. The financial perspective of the Balanced Scorecard—that all other perspectives build up to—does represent the bottom-line for an organization and a well-constructed Balanced Scorecard for a profit-maximizing entity should reflect profitability as its over-riding concern. What the other perspectives of the Balanced Scorecard do is to make more explicit and transparent the connection between profitability, on the one hand, and resources, operations, and customers, on the other.

The proponents of the Balanced Scorecard prefer to tackle the technical issues (e.g. metrics selection etc.) as pointed out by Meyer, through a "learning-by-doing" approach. Kaplan and Norton would argue that an organization needs to experiment with a family of measures for six months or a year before it can be determined whether the right set of

[50] Meyer (2002, pp. 6-7) suggests six requirements of an ideal performance measurement system, namely, parsimony, predictive ability, pervasiveness, stability, and applicability. The Balanced Scorecard falls short, it asserts, because it only caters for parsimony and predictive ability and fails to account for the rest of the four requirements.

[51] Meyer (2002, p.2)

measures have been selected. Kaplan and Norton (2001) cites examples of organizations that have used statistical data analysis to understand the correlation between performance measures and fine-tune their Balanced Scorecards—a possibility that Meyer (2002) acknowledges but insists may not be available to smaller organizations due to sample size limitations.

To be fair, Meyer's (2002) observations on the unintended consequences of using performance measures and of running down and gaming of performance measurement systems are real-life phenomenon whose importance cannot be underestimated. However, unlike Meyer (2002) that seems to take a very pessimistic view of humans' interaction with performance measurement activity, one can argue that individuals, normally well-intentioned, perform somewhere in between the two extremes. There is, nonetheless, much room for developing the technical foundations of the Balanced Scorecard and Meyer (2002) presents some excellent suggestions to that effect.

Austin (1996) also indirectly critiques the "balance" in the Balanced Scorecard as it asserts that today's performance measurement approaches, Balanced Scorecard included, "do not mention the importance… of measuring without missing any critical dimension of performance"[52]. It asserts that measuring only the easy-to-identify or easy-to-measure areas, especially at the expense of the critical-but-difficult-to-measure areas, is a flawed practice that creates the possibility that individuals would channel their productive energies towards those areas that are measured and fail to do what is critical for the organization.

Austin's critique can be construed as a valid one if one entertains the possibility, as Austin does, that a Balanced Scorecard cannot possibly encompass "all" critical dimensions of performance. Austin's approach to identifying the solution to this quagmire—based on traditional principal-agent models—is technically sound and thus deserves its due attention[53]. Austin's overall conclusion, however, is not that it cannot be done, but rather that organizational performance measurement is hard and one must be prepared, intellectually as well as from a management standpoint, to navigate through a difficult terrain. This is a valid observation and is not out-of-sync with the more thoughtful implementations of the Balanced Scorecards.

---

[52] Austin (1996, p.40)

[53] It is beyond the scope of this work to describe Austin's work in detail. We would, however, strongly recommend this work for the more thorough reader or anyone seeking a realistic assessment of the challenges involved in designing and implementing effective performance measurement systems.

3.3.2—Generic Challenges to the Balanced Scorecard

In addition to the above critiques and challenges, it would be an appropriate moment to consider a few interesting ideas that have substantive bearing on the validity of the Balanced Scorecard and the direction of the proposed research. One such set of ideas relate to the convergent and discriminant validity of the Balanced Scorecard. In other words, "Is there such a thing as a Balanced Scorecard?" The dual notions of convergent and discriminant validity, as adapted from Campbell and Fiske (1959) and applied to organizational context by Hackman and Wageman (1995), would assess the degree to which the Balanced Scorecard reflects a set of unique and consistent organizational interventions, structures, and systems designed to achieve a pre-defined purpose (i.e. performance improvement). In this context, convergent validity would reflect the degree to which the different versions of the Balanced Scorecard—i.e. those promulgated by its founders, other management theorists and consultants, and those observed in organizational practice—share a common conceptual core. The discriminant validity, on the other hand, would refer to the degree to which the Balanced Scorecard philosophy and practice can be reliably distinguished from other models of organizational improvement, such as the quality models (TQM), Six Sigma, and management-by-objectives (MBO) etc.

The issue of convergent validity is quite evident from our discussion in section 3.2. We conclude that the Balanced Scorecard concept does pass the test of convergent validity, as adapted from Hackman and Wageman (1995). While there are significant challenges to the convergent validity of the Balanced Scorecard concept, as prescribed by its founders Kaplan and Norton, the basic insight (i.e. performance multi-dimensionality) and related structural features (as defined in section 3.2) have been applied more less consistently throughout the organizational landscape. One of the reasons for this is the strong influence of the movement's founders that has shaped and led the intellectual development of the basic concept. The commonalities between the Balanced Scorecard prescribed by the founders (Kaplan and Norton,1992, 1993, 1996, 2001, 2004) and other management theorists and consultants (e.g. Olve et al., 1999; Brown, 1996; Nevin, 2001 etc.) form convincing evidence in support of a common conceptual core around which organizational practice has developed.

Much less comforting, though, are challenges mounted by group(s) of theorists who have either adopted the term Balanced Scorecard (or simply Scorecard) to describe significantly different management approaches (e.g. Gupta, 2004; Rampersard, 2003) or have put forth overlapping (similar yet different) approaches preceding or in response to the Balanced Scorecard (e.g. Neely, 2003; Adams and Roberts, 1993; or McNair et al., 1990). Of the two types, the second type of challenges are the most threatening to the construct validity of the

Balanced Scorecard, with the first one being somewhat irrelevant to it. We, however, believe that the Balanced Scorecard successfully navigates through this test, simply because the relative popularity of Kaplan and Norton's Balanced Scorecard dwarfs the influence of the competing frameworks.

FIGURE 3.9: TQM'S PHILOSOPHY, ASSUMPTIONS, PRINCIPLES, AND INTERVENTIONS[54]

- **PHILOSOPHY & ASSUMPTIONS**
  - The cost of poor quality is higher than the cost of processes to develop quality outputs
  - Employees care about quality work and, provided opportunities, will improve it
  - Organizations are interdependent systems and their problems are cross-functional
  - Quality is ultimately and inescapably the responsibility of top-management
- **CHANGE PRINCIPLES**
  - Focus on work processes—not outcomes alone
  - Focus on analysis of variability. Uncontrolled variance is primary cause of quality issues
  - Manage by Fact. Systematically and scientifically collect data to aid management decision.
  - Focus on learning & continuous improvement. Quality is a never-ending quest.
- **INTERVENTIONS**
  - Customer is the boss. Explicit identification and measurement of customer requirements.
  - Quest for quality begins at vendors. Creation of supplier qualification and partnerships.
  - Use scientific methods to monitor performance and identify points of high leverage.
    - Control Charts
    - Pareto Analysis
    - Cost of Quality Analysis
  - Manage Process. Use process management heuristics to enhance effectiveness
    - Flow charts (to pictorially represent work processes)
    - Brainstorming (to generate ideas, analyses, solutions, and issues)
    - Fishbone diagrams (to represent cause-and-effect linkages)

The issue of discriminant validity presents a more challenging situation. While there does exist a distinct conceptual core of the Balanced Scorecard, as described by its structural and derived features (section 3.2), it also has significant overlaps with other management fads of the past. This raises the question of whether much of what goes on in the name of the Balanced Scorecard is essentially a re-packaging of popular managerial practices that have been proposed over time. For example, Hackman and Wageman (1995) present an in-depth analysis of TQM's philosophy, its underlying assumptions, its change principles, and interventions. We summarize this information in Figure 3.9 and use it to address the issue of discriminant validity of the Balanced Scorecard.

[54] This table draws heavily from Hackman and Wageman (1995)

Looking at the menu of constructs and principles identified in the Figure 3.9, the similarities and differences between TQM and the Balanced Scorecard become quite apparent. There are significant differences in the basic philosophy behind TQM and the Balanced Scorecard with the former focusing solely on quality as over-riding concern of organizational leaders while the latter taking a more "balanced" view of organizational performance. Both approaches focus on work processes to improve organizational performance, albeit to varying degrees. Internal business process is just one of the several perspectives on a Balanced Scorecard but it is the central core of TQM. Both frameworks emphasize on "management by facts (or measurement)" although TQM is much more scientific (at least in theory, if not practice) about measurement of performance than the Balanced Scorecard. Both frameworks depend on learning and continuous improvement through employee initiatives.

At a broad level, they also have several interventions in common including, but not limited to, pictorial representations of work processes, cause-and-effect diagrams, and participative brainstorming sessions. In the details, however, there are significant differences. The Balanced Scorecard uses strategy-maps as a pictorial aid to organizational communication. Similarly, the cause-and-effect modeling in the Balanced Scorecard is qualitatively different from fishbone diagrams and work-flow diagrams. The focus of the Balanced Scorecard's strategy maps is the organizational strategy while that for TQM is the process itself. These differences—"measuring the strategy" vs. "measuring the process"—are quite fundamental in nature thus strengthening the case for discriminant validity of the Balanced Scorecard.

One potential factor, however, that weakens this case is the tendency of the Balanced Scorecard to adopt organizational improvement interventions that form key parts of other such management frameworks and approaches. This is, in part, a conscious strategy one the part of the movement's founders of presenting the Balanced Scorecard as an integrating framework for organization's performance improvement and strategic management initiatives (Norton, 2002,) perhaps, as a means to counter the challenge of obsolescence. On the whole, however, it would be safe to suggest that the Balanced Scorecard does reflect significant fundamental differences from other organizational improvement approaches to pass the discriminant validity test. It does so much more convincingly for its conceptual core than for the structural-operational paraphernalia.

Having defined the conceptual core and structural paraphernalia of the Balanced Scorecard and addressed some of the methodological and philosophical challenges to it, we would now turn to adapting the concept to research and development organizations (RDOs).

3.4—APPLYING THE BALANCED SCORECARD TO R&D

While the use of the Balanced Scorecard has pervaded much of the corporate world as as well healthcare, IT-service delivery, non-profits, public sector agencies, and even city and country governments[55] its use has been rather limited within the R&D[56]. This is a trend that is beginning to change now.

As discussed earlier, the notion of "balance" is not new to the R&D world (e.g. Brown & Svenson [1988] have suggested the idea that a "balance" be achieved between input, process, output and outcome measures of an R&D system,) The earliest attempt to replicate a Balanced Scorecard within R&D came in the form of "Technology Value Pyramid" (Tipping et al., 1995.) In many ways, the Technology Value Pyramid (TVP) was an R&D-equivalent of a Balanced Scorecard—officially championed and patronized by the Industrial Research Institute (IRI)—the official body of Industrial R&D Labs in the United States.

TVP divides the performance dimensions of an R&D organization into three categories, namely, outcome, strategy, and foundations. These three performance categories (or dimensions) are further subdivided into five managerial factors (or sub-categories), namely, value creation (in outcomes category), portfolio assessment and integration with business (in strategy category), asset value of technology and practice of R&D processes (in foundation category). By picking a set of metrics—from a menu of metrics[57]—within each of these five sub-categories of performance—TVP ensured balance across various performance dimensions. By standardizing the performance dimensions themselves, TVP ensured that a chain of causality could be established from the foundation (i.e. the R&D processes within an organization) right up to the desired outcomes (i.e. value created by R&D.)

Although there are significant similarities between TVP and the Balanced Scorecard, especially in terms of the organizing structure (i.e. aspiration to achieve balance across multiple performance dimensions, and cause-and-effect linkages that permeate the entire value chain of R&D within an organization)—there are major differences as well. The most

[55] The most notable of the use of the Balanced Scorecard at the level of city and country government have been the City of Charlotte in North Carolina and President Vincent Fox's Mexico respectively. (Kaplan & Norton, 2001, 2002 etc.)

[56] Personal conversation with Dr. Robert S. Kaplan, Boston, December 2002.

[57] A menu of 33 metrics was developed after 2 years of consultations between members of the Research-on-Research sub-Committee of the Industrial Research Institute (IRI) and its Quality Directors Network.

significant of these differences are procedural rather than structural. TVP does not pay much emphasis on the implementation process. In fact, Tipping et al. (1995, p.39) suggested the use of an expert system-based program to allow users of the TVP to easily select an appropriate set of measures for their particular circumstances (i.e. type of organization, and R&D activity.) Another possible shortcoming of TVP was its rather weak linkage with organizational strategy—at least not in a manner that is a characteristic of the Balanced Scorecard.

Given that the Technology Value Pyramid was motivated by the Balanced Scorecard itself[58], it is interesting to take note of the fact that its founders, unlike Kaplan and Norton, chose not to emphasize the organizational strategy and communication aspects of the model. They were also very modest in terms of their claims vis-à-vis TVP's ability to influence performance. In fact, they were careful enough to isolate the differences between performance measurement and performance management functions. To that effect, they stressed that the "use of the model, in and of itself, will not improve an organization; measurements only provide data…improvement depends on R&D management action consciously taken to improve various measures of performance within the selected set of measurements" (Tipping et al., 1995, p. 39.)

Since Tipping et al. (1995,), we have seen a growing sense of awareness about the Balanced Scorecard among the R&D management community—as evidenced by publication of articles in professional engineering and R&D management journals (e.g. Crandall, 2002; Kerssens-van Drongelen et al., 1997 and 1999; Loch and Tapper, 2002; and Bremser and Barsky, 2004) as well as instances of actual implementation of the framework in real-life R&D organizations.

Kerssens-van Drongelen and Cook (1997)—a theoretical article on the design principles of performance measurement systems for R&D processes—talks about the appropriateness of the Balanced Scorecard for R&D organizations. It provides a snapshot of a straight-forward adaptation of a corporate Balanced Scorecard to the R&D environment. This is reproduced in Figure 3.10 (below). Kerssens-van Drongelen et al. (1999) puts this model to an empirical test. A survey of 225 R&D managers in the Netherlands was conducted of which 48 (21%) responded. Of the respondents about 80% measured R&D in one way or the other. The respondents were asked about the performance metrics they used to measure the performance of their organization at the individual, team, department, and company level. The authors tried to list the metrics thus collected in the four Balanced Scorecard

[58] Phone Conversation, Alan Fusfeld, October, 2003

performance perspectives in an effort to assess how well do actual R&D organizations fare in terms of using a balanced set of performance measures. The results are somewhat revealing. We would only discuss the results at the level of the company. The more

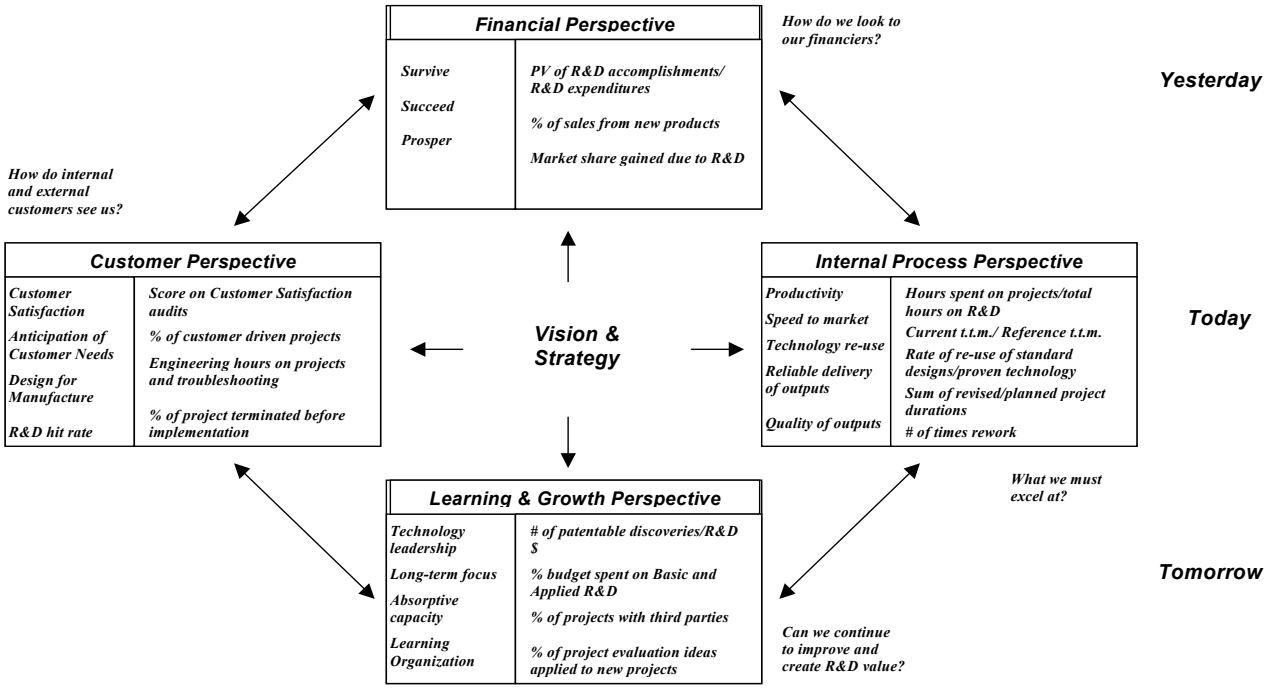


Figure 3.10: Example of a Balanced Scorecard for an R&D Organization (Source: Kerssens-van Drongelen et al., 1997)

inquisitive reader is advised to look at the study for detailed results.

At the level of the overall company, only 33% of the respondents measured customer satisfaction or market response of their R&D products. The most highly measured dimension was the internal business process dimension where 50% of the organizations measured number of projects completed and another 28% measured speed of product development. Efficiency and quality of work (11% each) were other two metrics in this category. Innovation and learning perspective was the least measured of the four performance dimensions. Only 11% of the companies measured number of patents, and another 6% measured number of ideas. Around 80% of companies seemed to not report any measure in the learning and innovation category at all. The two key metrics within the financial performance perspective included new sales ratio (28%) and profit due to R&D (22%) with only 11% reporting measuring expected or realized ROI from R&D. The results are instructive as they are somewhat indicative of the degree of balance (or imbalance)

across organizations' R&D measurement activities, but they are also somewhat simplistic as they only relate to a very narrow interpretation of the Balanced Scorecard.

Bremser and Barsky (2004) attempts to take up from where Kerssens-van Drongelen et al. (1997, 1999) have left and build upon it. In search of an integrated performance measurement architecture for R&D, they attempt to use Kaplan and Norton's Balanced Scorecard in conjunction with the stage-gate approach commonly used in R&D organizations. The key challenge in measuring performance of R&D organizations, per Bremser and Barsky, is to use past-focused cost data with prospective long-term strategic and financial objectives—something a Balanced Scorecard's multi-attribute structure is well-placed to do. While the authors acknowledge Kerssens-van Drongelen et al.'s (1997, 1999) contribution to the literature, they criticize it for failing to show whether any of the firms surveyed had actually implemented a Balanced Scorecard in R&D. The paper itself extends the literature only add to the literature in discussing how measures obtained from the stage-gate process could be used to extend the firm-level metrics of an R&D Balanced Scorecard to the next level of the R&D organization's hierarchy, namely, the R&D department. The article does not address the question of whether and how do R&D organizations implement a Balanced Scorecard or benefits that might accrue as a result of it.

Papers in the tradition of Kerssens-van Drongelen et al. (1997, 1999 etc.) and Bremser and Barsky (2004) represent, to a great extent, the state-of-the-art in the literature that has only begun to explore the applicability of Balanced Scorecard-type methodologies R&D. Clearly, this is just the beginning and this relatively small and growing literature raises more questions than it is able to address. First, and most importantly, as pointed out by Bremser and Barksy (2004), while Kerssens-van Drongelen et al. (1997, 1999) talk about balance across multiple dimensions of performance, it does not state whether any of the firms surveyed actually used the Balanced Scorecard to measure and manage performance. Both these papers also do not look at the process dimension of implementing a Balanced Scorecard, as well as the presence (or lack of) other important structural features (e.g. cause-and-effect logic, link with strategy and strategy-mapping.)

These papers are also far from addressing the more philosophical issues about the use of the Balanced Scorecard in R&D, namely: are R&D organizations different, in any substantive way, from the corporate sector organizations for which the Balanced Scorecard was initially developed? Do we need a modified version of a Balanced Scorecard to go with R&D organizations? We address these questions in the following section. The absence of literature on actual use of the Balanced Scorecard within R&D, with the exception of Loch

and Tapper (2002), presents a picture of a completely uncharted terrain. It is against this backdrop, and to address the above questions, that this research was proposed.

3.5—TOWARD A GENERIC R&D BALANCED SCORECARD

An attempt to assess the appropriateness of the Balanced Scorecard for an R&D organization must begin at deciding upon a structure of a generic R&D Balanced Scorecard. As clearly evident from the discussion on structural and derived features, the Balanced Scorecard hinges on a number of structural and process elements whose amenability to the R&D environment must also be individually assessed. The next chapter would draw an analytical agenda and approach to address some of these questions. In this section, however, we present a suggested structure of a generic R&D Balanced Scorecard—drawing upon the vast literature in R&D management and measurement and the Balanced Scorecard.

3.5.1—The Critical Dimensions of R&D Performance (the Perspectives)

The most important and defining structural feature of the Balanced Scorecard is the notion of performance dimensions (or perspectives). As discussed above, the traditional Balanced Scorecard uses four causally-linked dimensions of performance for an organization. Brown (1996) and Olve et al. (1999) discuss the possibility of adding a fifth, namely, the human resources dimension in order to stress upon the importance of developing, valuing, and measuring the state of human resources in an organization. Kaplan and Norton (2001) also allows for flexibility in the structure of the Balanced Scorecard e.g. renaming performance dimensions, adding categories, and re-structuring the relationship between categories to reflect the reality of the organization in question. One of the legitimate reasons for doing so, we believe, is to make the concept of the Balanced Scorecard more amenable to the type of audience one is dealing with. In the context of the current exercise, for example, this would mean using terms, metrics, and measurement approaches that R&D employees (i.e. scientists and engineers) can relate to and feel comfortable with. Consequently, our R&D Balanced Scorecard and the generic Strategy Map contains five dimensions of performance. These include:

- ***Employee morale and creativity (EMC) dimension.*** Employees are perhaps the most valuable assets of any organization, and this is especially true for an R&D organization where employee morale and creativity drives output and performance. Maintaining a high-quality workforce takes considerable management attention and requires the right kinds of systems, structures and processes. These could include recruiting and hiring systems, performance measurement systems, reward and

recognition systems, career progression models, and employee education and training. Measuring the performance of these systems (and sub-systems) as well as the overall state of employee morale is critical to the overall health of an R&D organization.

- ***Innovation management (IM) dimension.*** Innovation management is akin to the internal business process management perspective of a traditional Balanced Scorecard. R&D organizations create value through innovation. They rely on specific systems, structures, processes, and initiatives to create knowledge and develop new technologies or products. Examples include strategy formulation processes, requirements generation processes, systems that determine techno-market risk of technical projects (e.g. the stage-gate process), and project management systems. The measures of the efficiency and effectiveness of these internal processes of an R&D organization constitute the process dimension of the R&D Balanced Scorecard.

- ***Organizational learning, dissemination, and knowledge management (LKM) dimension.*** Organizational learning, research dissemination, and knowledge management are critical to the health of an R&D organization. Science is a social enterprise (Hurley, 1997) that depends on the fruitfulness of these interactions within organizations and communities. Organizations that do a good job of capitalizing on their in-house knowledge and capabilities, while making efficient use of external capabilities and resources, generally perform better than those that do not. World-class research organizations use a number of systems, structures, and mechanisms to ensure continuous learning.[59] Preservation of corporate memory is another important element of knowledge management, particularly in an era of high employee turnover. Finally, the use of systems and practices to maximize "learning from one's mistakes" is a key facet of organizational learning (Branscomb et al. 2001). Measuring how (and how well) organizations learn and utilize their knowledge-base is another critical performance dimension of R&D performance.

- ***Financial control and performance (FPC) dimension.*** While the financial perspective represents the desired end-result for a traditional Balanced Scorecard—the

[59] These include exchange programs with other parts of the knowledge/innovation system (e.g. university-industry exchanges, entrepreneurial sabbaticals), participation in standard professional dissemination efforts (e.g. industry conferences, workshops, colloquia and exhibitions), incentivizing their scientists and engineers to contribute to journals and professional meetings in their respective area and draw from it, as well as development of formal knowledge and competency management systems.

proverbial bottom-line, so to speak—this is often not the case in R&D. The bottom-line for an R&D organization is measured in terms of its innovative output or contribution to corporate parent or financial sponsor. This, however, should not create an impression that financial performance is not important for R&D organizations. To be fair, financial considerations often loom large among the issues that R&D managers are deeply concerned about. They still have to raise money for R&D and ensure an efficient utilization of these resources, justify their existence to their corporate parents or financial sponsors, and increasingly "do more with less". Measuring the financial performance is thus critical to assessing the overall performance of an R&D organization—for acquiring and managing money wisely makes everything else possible—even though it may not be the all-important bottom-line.

- ***Customer satisfaction (CS) dimension***. Like any organization, R&D organizations have internal or external customers whose satisfaction is vital to the organization's long-term health. These customers can either be internal business units (e.g. in the case of a development lab,) corporate top management (e.g. a in the case of a central lab,) government agencies (e.g. in the case of DOE's National Laboratory system, NASA, or DOD laboratories,) or the society and the intellectual community at large (e.g. in the case of publicly-funded R&D institutions like NSF and NIH, and university-based academic research.) In each of these cases, the R&D lab's management is responsible for ensuring that the customer is satisfied and thus willing to provide continued support. Given the importance of customer satisfaction, it is vital that systems, structures, and processes be put in place to ensure that the organization measures and monitors customer satisfaction, maintains its ability to sense customer needs, and responds to them in a timely fashion.

These five performance dimensions, we believe, adequately reflect the critical dimensions of performance in an inclusive yet parsimonious manner. Why these five perspectives, and not another sixth or an entirely different list at all? The easiest answer to this critique would be to consider the above five performance dimensions as hypothesized. Specifically, we do not, at this point, claim to suggest that only these five dimensions of performance matter to R&D organizations. Instead, our approach is much more exploratory and conceptual. Allowing for individual diversity, we hypothesize these to be the most important dimensions of performance that can be generalized across a variety of R&D organizations. We would put this hypothesis to a test to ascertain its validity across various types of R&D organizations.

3.5.2—The Relationship Between Performance Dimensions (the "Architecture")

Having established the critical dimensions of performance in R&D, we now turn toward the relationship between them as formalized in the architecture of the Balanced Scorecard's strategy maps. . The panel on the left (Figure 3.11, below) presents one of the several alternatives. This kind of architecture[60] has been sometimes used within the non-profit or public sector where customer rather than financial concerns have dominated the organizations' performance agenda (Nevin, 2003). It is based on the notion that all performance dimensions are interlinked with financial performance being the foundation of the overall organizational performance and customer satisfaction forming the desired end-result. In this scheme, the Innovation Management dimension is of central importance as it determines the core of organizational performance, namely, the efficiency of its work processes, with employee and learning and knowledge dimensions supporting it.

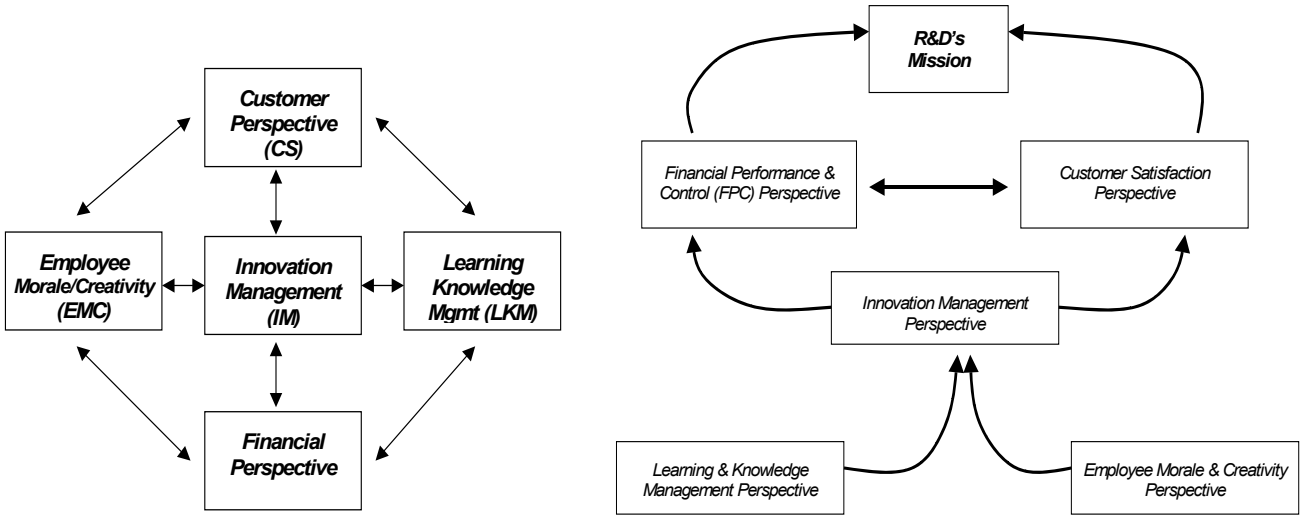


Figure 3.11: Dimensions of R&D Performance—Two Alternate Architectures

An alternate architecture is presented in the panel on the right (Figure 3.11, above). This uses an alternate logic of relationships between various performance dimensions. This logic suggests that Employees Morale and Creativity (EMC) and the Learning and Knowledge Management (LKM) dimensions form the foundations on which an R&D operation is built. These lead to efficient and effective performance along the Innovation Management (IM) dimension. The key difference between the two architectures is the treatment of Financial Performance and Control (FPC). While the former treats the financial perspective as the foundation on which the R&D organization is built, the latter considers it as one of the two

[60] Nevin (2003) calls this schematic, the geography of the Balanced Scorecard.

end-results of its operation—the other being customer satisfaction. In this view, the customer satisfaction and the financial performance control perspectives are intertwined and overlapping. Several aspects of customer satisfaction (e.g. on-cost delivery, quality) have important financial ramifications. Also, many times an effort is made to gauge the end-result of the R&D process in financial terms (e.g. use of new product sales metric, or the internal rate of return from R&D.) Together, the customer satisfaction and the financial performance dimensions lead to the fulfillment of R&D's mission, namely, improve the competitiveness of its corporate sponsor through innovation, enhance knowledge available in the public domain, or advance the specific mission of a public agency. For the purpose of our generic strategy map for R&D organizations, we use the latter of the two architectures.

3.5.3—Putting it All Together—A Generic Strategy Map for R&D Organizations

The performance perspectives and their cause-and-effect interrelationships come together into the form of a strategy map. Strategy maps are graphical representations of an organization's differentiated strategy. It makes prominent the key actions and initiatives in the critical dimensions of an organizations performance and identifies performance metrics needed to measure the progress on each. In other words, it "tells the story" of the organization's strategy.

Figure 3.12 (below) presents a "generic" strategy map for an R&D Balanced Scorecard. It is generic because it represents an exhaustive list of broad-level features of an R&D organization—several of which may not be important or relevant to a particular organization at all but may be helpful in generating a dialogue about its strategy. In fact, some of these features are mutually exclusive to each other.

Like all strategy maps found in the Balanced Scorecard literature (Kaplan and Norton, 2001 and 2004; Nevin, 2003 etc.), the generic strategy map for an R&D organization begins with the ultimate objective—the reason for being—of the organization, namely, its mission. We divide the mission of an R&D organization into two components, namely, the stated and the unstated. The stated mission of an R&D organization is what it is charged with, explicitly, by its direct and primary customers. The unstated mission of an R&D organization, one the other hand, refers to the needs of its indirect customers or even the unstated needs of its direct customers.

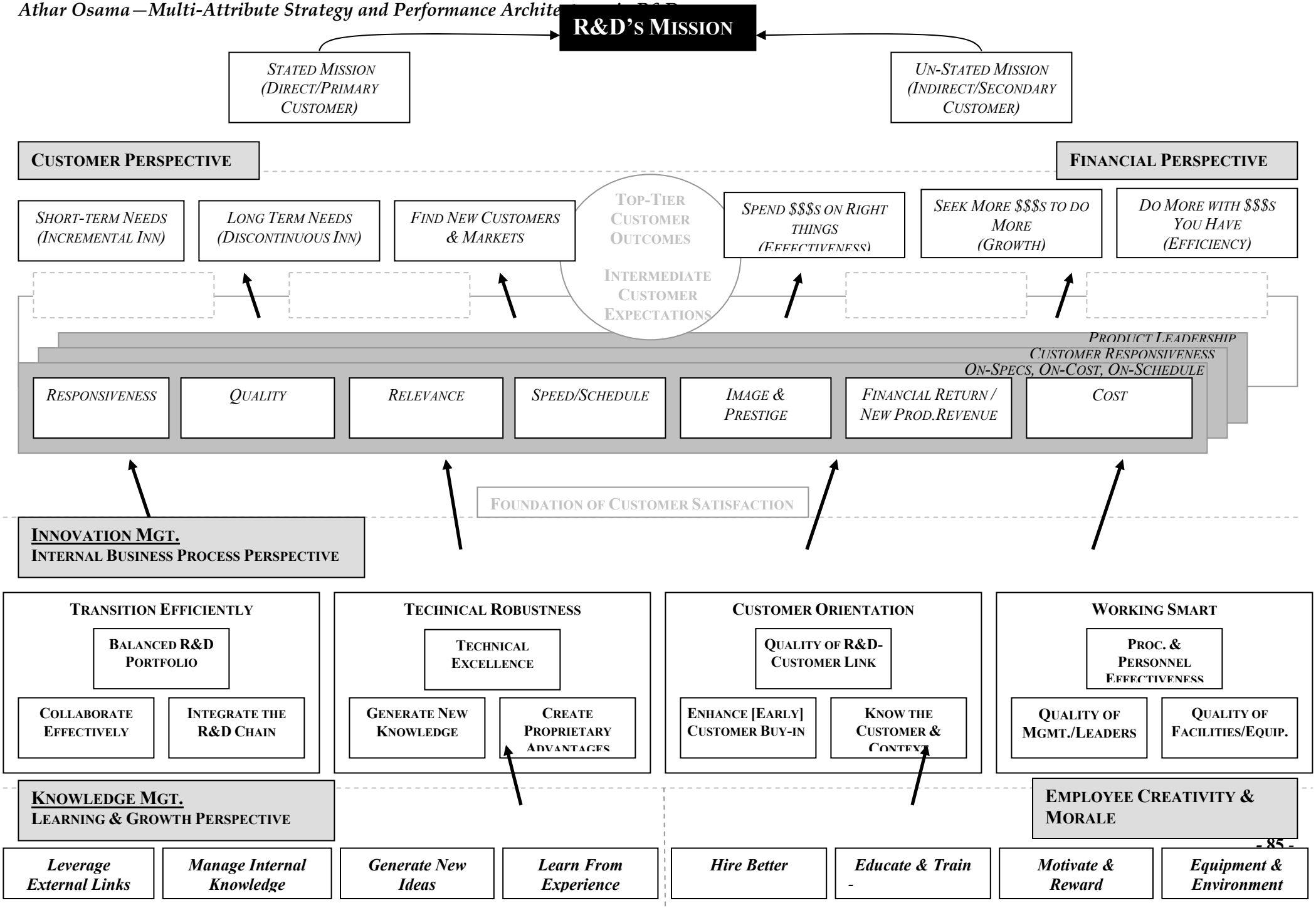


Figure 3.12: A Generic Strategy Map for an R&D Balanced Scorecard

This establishes the focus toward the satisfaction of both direct and indirect customers. There are three broad-level generic themes ("top-tier" customer outcomes) in the customer perspective, namely:

- ***Meeting the short-term needs*** (refers to incremental innovation),
- ***Meeting the long-term needs*** (refers to discontinuous innovation), and
- ***Finding new customers and markets*** for its R&D services.

These broad-level themes, however, are just that. To look at the construct of customer satisfaction at a deeper level, one must de-construct the determinants of customer satisfaction. We identify seven of these on the generic strategy map, namely, responsiveness, quality, relevance, speed or schedule, image, financial return, and cost. This is clearly a partial and suggestive list of attributes that form the foundation of customer satisfaction in R&D. The final determination of which of these factors matter to a specific organization's customers is remains a matter for the judgment of an organization's leaders. We also accommodate the possibility of identifying a layer of mid-level ("intermediate") measures that elucidate the relationship between the foundational factors and the top-tier outcomes. These can be tailored to the specific needs of a particular organization in question.

The financial performance and control perspective overlaps with the customer perspective in significant ways. Customer satisfaction is a multi-dimensional attribute that depends, among others, on cost and return on investment considerations as well. The financial perspective also displays three broad-level generic themes, namely:

- ***Efficiency***—doing more research and development with less resources;
- ***Effectiveness***—doing R&D of the right kind with given resources; and
- ***Budgetary Growth***—finding more resources to do research and development.

This is also a good opportunity to introduce the notion of customer value propositions, as described by Michael Treacy and Fred Wiersema in *The Discipline of Market Leaders* (Perseus Books, 1995), and adopted by Kaplan and Norton (1996, 2001). According to this view, organizations need to adopt one of the three core customer strategies, namely, product leadership, customer intimacy, and operational excellence—and never all of them at the same time. Strategy-focused organizations, it argues, do not try to be everything for everyone. Porter (1996) suggests that strategy is about making choices. Do R&D organizations adopt this differentiated view of strategy or do they try to be everything for everyone, remains, as of now, an intellectual curiosity.

We incorporate the idea of core customer strategies in our generic R&D strategy-map to allow R&D organizations to adopt differentiated strategies. These are technical excellence, operational excellence, and customer responsiveness. What we expect to see, if R&D organizations take a differential view of strategy, is for organizations to pay greater emphasis on those aspects of customer satisfaction outcomes that are related to their expressed strategy. For example, if an R&D organization's expressed objective is to do research at the cutting-edge of its area of expertise (i.e. product leadership or technical excellence,) it must pay greater emphasis on customer satisfaction attributes related to technical excellence (e.g. quality and innovativeness of its work) rather than cost or schedule.

In fact, not only would an R&D organization emphasize on customer satisfaction attributes relevant to its expressed objective, it would also explicitly develop its innovation and innovation management processes to support that. We identify four broad themes within the innovation management perspective of the generic R&D strategy-map. Three of these themes relate to the three core customer strategies identified above. These include:

- ***Working Smart*** relates to the customer strategy of by the operational excellence.
- ***Customer Orientation*** relates to the customer strategy of customer responsiveness.
- ***Technical Robustness*** relates to the customer strategy of technical excellence.

A fourth theme of the innovation management perspective is technology-transition. We included this because of its importance to the innovation processes of all types of R&D organizations and expect it to be common across all.

Finally, the last two perspectives, namely, employee morale and creativity and learning and knowledge management form the foundation of the generic R&D strategy-map. These activities also form the foundation on which an R&D organization is built. The general ideas included in the employee morale and creativity perspective relate to hiring better employees, training them well to meet the challenges of research, motivating employees to perform, and providing them with the right equipment and environment. R&D organizations might choose to undertake these tasks in different ways. Having an explicit strategy—for and keeping a track of, how ones organization is performing in all of the above respects—is important part of R&D management. Similarly, in the learning and knowledge management perspective, we included activities like optimal use of internal and external knowledge, learning from experience, and developing and capitalizing on new ideas. Different organizations may choose to implement these ideas differently and these

choices may be reflected as the choice of various initiatives or performance measures on an organization's Balanced Scorecard.

Alternatively, organizations may choose an entirely different set of ideas, unique to their own strategy, under each of the performance perspectives in which case the generic R&D scorecard may be used as a means to start a strategy dialogue and would ultimately be modified to reflect the reality of that particular organization.

3.5.4—Doing It Right (the "Process")

Establishing the structure and architecture that makes sense is merely a start. Implementing a Balanced Scorecard successfully, however, also requires putting in place the right processes. Managing the process aspects of the Balanced Scorecard implementation like participative strategy-making, intra-organizational communication, cascading, cause-and-effect thinking, and strategic learning etc. are critical without which any implementation effort would be at best incomplete, and at worst, faulty and misguided.

Therefore, an R&D organization seeking to implement a Balanced Scorecard must carefully internalize and incorporate the process dimension of the Balanced Scorecard, in addition to is structural elements. The organization's leadership must also consider the possibility that the nature of the R&D environment or the workforce might require making some changes in the process to make it more amenable to R&D. For example, participative strategy-making, as prescribed by Kaplan and Norton (2001) may or may not be readily acceptable or even the most appropriate practice within the context of the R&D organizations. With no documented evidence of actual practice to guide us, we are left to speculate about what would constitute a good process for implementing a Balanced Scorecard within an R&D environment.

We started this chapter with putting together a frame of reference for talking about the Balanced Scorecard. We developed a benchmark model of the Balanced Scorecard and discussed the various structural and process features that make up well-formulated and implemented Balanced Scorecards within private sector. We then used this benchmark model to develop a generic Balanced Scorecard for an R&D organization. Thus far, however, our generic model of an R&D organization is as good as anybody else's. The logical extension of the work done so far is to assess the validity of our generic framework against actual practice. How prevalent is the use of the Balanced Scorecard among R&D organizations? Do R&D organizations take a multi-dimensional view of strategy and performance? Do R&D organizations take a differentiated view of strategy or do they try to be "everything for everybody"? Do R&D organizations that seek to implement the Balanced

Scorecard adopt a modified version or use the traditional four-dimensional structure? What process elements do R&D organizations implementing a Balanced Scorecard emphasize upon? What results do they get from their efforts? These questions cannot be addressed and validated without carefully studying a cross section of R&D organizations.

In the chapter that follows, we develop an analytical agenda to address some of the questions identified above and describe an approach to doing that.

CHAPTER—4

THE RESEARCH QUESTIONS & METHODOLOGY

Thus far, we have developed an analytic framework that details the various structural and derived features of the traditional Balanced Scorecard, as it has been applied to a variety of organizational settings. We also developed a checklist against which efforts to create and implement Balanced Scorecard initiatives may be compared to assess their completeness and quality. Finally, we developed a generic Balanced Scorecard and Strategy Map for R&D organizations. The task at hand now is to draw upon the above frameworks and compare and validate these against actual practice on the ground and thus make some conclusions and recommendations about the appropriateness of using the Balanced Scorecard methodology in R&D settings This calls for the definition of an analytical agenda—a set of research questions and an analytic approach to answer these—a task that we would undertake in this chapter. Section 4.1 describes the analytical agenda for the above-mentioned project and Section 4.2 presents a methodology to address it.

4.1—THE POLICY & RESEARCH QUESTIONS

The proposed research would address questions at two levels, a broader philosophical level, and a narrower and more pragmatic level. While the main focus of our investigation is the narrower and pragmatic level that addresses the questions that arise from the application of the Balanced Scorecard in R&D, gleaning from the broader philosophical analysis sets the stage of interpreting the narrower set of recommendations by putting these in the proper context. We describe each of these in some detail below:

At the broader and more exploratory level, the study aims to explore ideas and constructs relating to the prevalence and nature of performance measurement in research and development organizations and the differences in performance measurement philosophy and methodologies across various types of research activities (e.g. basic, applied, development etc.) and organizations (e.g. public sector labs, corporate labs, and academia etc.). These questions include:

- ***G.1—How do various types of research and development organizations (RDOs) think about and measure their performance?*** Measurement, as discussed in chapter-2, has long been a controversial idea within the R&D community. While this may be changing over time, how organizational leaders, managers, and bench scientists and engineers think about measurement is important for performance measurement system design. In

the context of the current research, we are particularly interested in assessing how perceptions about performance measurement in R&D vary across different types of R&D organizations and activities. More specifically:

  - G.1.1—How do R&D leaders and managers think about performance measurement in R&D?
  - G.1.2—What do R&D managers believe are the legitimate uses of performance information in R&D and what are some of the commonly prevalent ways organizations put their performance information to use? How does that differ across type of R&D work performed and organizational arrangement used to perform that work?
  - G.1.3—What commonly prevalent measurement philosophies, methodologies, and tools as well as metrics do they use? What do these choices tell us about the underlying philosophy of their measurement efforts?

- ***G.2—Do R&D organizations take a multi-dimensional view of organizational performance and do their institute performance measurement systems reflect that view?*** That organizational performance is a multi-dimensional construct is well-established in literature (Kaplan and Norton, 1992 etc.; Maisel 1992; McNair et al., 1990.) Effective performance measurement systems must incorporate performance multi-dimensionality and do so in several ways, simultaneously. Not only must they measure performance across the various stages of input, process, output, and outcomes but also across the perspectives of various stakeholders (e.g. employees, shareholders, customers,) and types of indicators (leading and lagging, qualitative and quantitative, and financial and non-financial) etc. At an intuitive level, R&D organizations have long entertained the notions of performance multi-dimensionality in stages of activity (e.g. Brown and Sevenson, 1988; Baglieri et al., 2001,) type of R&D (e.g. Hauser and Zettelmeyer, 1997,) and types of metrics (e.g. Pappas and Remer, 1985.) More recently, there has also been a move towards incorporating the perspectives of other stakeholders e.g. customers, sponsors, and employees in performance assessment schemes. We are specifically interested in:

  - G.2.1—What are the common dimensions of performance deemed critical by R&D managers? How do these performance dimensions vary across types of R&D and organizations?
  - G.2.2—Do R&D organizations that take a multi-dimensional view of performance tend to perform better than those that do not?

- ***G.3—To what extent do performance measurement systems in R&D organizations incorporate a set of generic structural features essential for developing effective R&D measurement systems?*** We use our generic framework for effective R&D organizations (developed in chapter-3) to assess the extent to which these implement a full-set of

structural features needed to exploit the potential of a performance measurement framework. Most performance measurement approaches come in bundles of individual components that are related and must feed into each other to bring about improvements in performance. The difficulties of assessing the extent of implementation of such management frameworks and their effects on organizations are well-documented in literature (e.g. Hackman and Wageman, 1995.) Several sub-components of the R&D measurement problem (e.g. metrics selection model) have attracted lesser analytic attention than others (e.g. metrics properties) thus leading to situations where organizations have failed to implement a full set of structural features necessary to drive improvements. We are interested in exploring the extent to which selective adoption of features for reasons of conscious choices of convenience or lack of appropriate guidance--is a norm rather than an exception. More specifically:

  - G.3.1—Do R&D performance measurement systems address each of the five considerations outlined in the measurement problem framework?
  - G.3.2—Do R&D performance measurement systems address the structural and derived features outlined in the generic analytic framework?
  - G.3.3—Do R&D organizations with more complete performance measurement systems (with completeness assessed by presence or absence of various structural and derived features) tend to perform better than those with incomplete ones?

- ***G.4—How do R&D organizations tie performance measurement systems with organizational strategy, on the one hand, and incentive design, on the other?*** The Balanced Scorecard literature firmly establishes the relationship between performance measurement and organizational strategy, namely, that the former must "measure[s] the strategy" (e.g. Kaplan and Norton, 1996, 2001.) Strategic alignment of R&D has long bothered R&D leaders and yet, with the exception of Tipping et al. (1995,) the field is devoid of well-developed management frameworks that tie R&D performance with organizational strategy. The measurement of organizational strategy has had its own set of problems (Ginsberg, 1988; Inkpen and Choudhary, 1995; Nath and Sudharshan, 1994). On the other hand, some have faulted the lack of a link between strategy, performance, and incentives system in organizations' ability to execute strategies (Stonich, 1981). We are interested in assessing how R&D organizations link their performance measurement systems to reflect organizational strategy and use incentives to encourage implementation. Specifically, we are interested in:

  - G.4.1—To what extent do R&D organizations measure what they claim to be trying to achieve (i.e. their objectives and strategy)?
  - G.4.2—How do R&D organizations link organizational and individual performance measurement with incentives structures?

  - G.4.3—What is the effect of the quality of the strategy-performance-incentives connection on the effectiveness of the R&D performance measurement system itself?

- ***G.5—Are there any differences between public sector, academic, and private sector R&D organizations in strategy formulation, performance measurement, and incentives design?*** Authors have long analyzed differences between private-sector, public-sector, and academic R&D. Many have contended that public sector organizations, by virtue of their mission and orientations and multiplicity of constituencies, are different and respond differently to organizational interventions. Others have contended that "public-ness" is not a discrete organizational characteristic that an organization is born with but rather a continuous variable that is the characteristic of all organizations (Bozeman, 1988.) There is also no dearth of literature on the unique organizational characteristics and motivations in the university and academic environments (e.g. Diamond, 1999; Brewer et al., 2002; Hauser and McEnaney, 2000.) We are interested in exploring the differences in strategy-formulation, performance measurement, and incentives environment in public and private-sectors and academic labs. More specifically, we are interested in:
  - G.5.1—What are the major differences in strategy-formulation, performance measurement, and incentives between public, private, and academic R&D settings?
  - G.5.2—Do these differences seem to arise from the differing nature of their work, organizational objectives, or their organizational structures?
  - G.5.3—Given the differences, how might the performance measurement systems be modified to best suit the needs of these various types of organizations?

As we approach and engage with the subject of this research more substantively, we attempt to develop a sense of the above described landscape. Being able to answer these questions would help us understand the overall environment in which the performance measurement function is put to work within R&D and highlight the differences between R&D and other types of business organizations. This would create the context and perspective in which a more focused study of the Balanced Scorecard in R&D could be undertaken. In line with the above objectives, we approach these questions in an exploratory rather that explanatory spirit without any prior hypotheses to falsify.

At the narrower and more pragmatic level, the proposed research project aims to assess the appropriateness of Balanced Scorecard-type multi-attribute performance measurement systems (MAPMS) in R&D. We are specifically interested in making generalizable conclusions and recommendations that might lead to the modification and/or

implementation of the Balanced Scorecard methodology to suit the unique requirements of the R&D organizations. This task would entail testing for the key (structural features and) assumptions that make the Balanced Scorecard an effective strategy-performance-incentives. We are seeking to answer a series of questions relating to the prevalence, nature, and effectiveness of Balanced Scorecard-type strategy-performance-incentives architectures in R&D. These include:

- ***HS.1—Prevalence of the Balanced Scorecard***— That Balanced Scorecard-type strategy-performance-incentives architectures are not as popular in our population of interest as they are in other commercial and non-commercial settings is a well-known fact (Kaplan and Norton, 2002; Fusfeld, 2003.) The R&D management and measurement literature mentions the Balanced Scorecard only indirectly (e.g. Kerssens-van Drongelen et al., 1999; Bremser and Barsky, 2004.) Formal studies of the actual implementation of Balanced Scorecard in R&D are few and far between (e.g. Loch and Tapper, 2003.) In the absence of any statistical evidence to support the contrary, one would hypothesize that Balanced Scorecard-type frameworks are not as prevalent in R&D as some other similar approaches (e.g. TQM, Peer Review, Six Sigma) are.

  Performance multi-dimensionality, however, is a well-established construct in the R&D community (e.g. Brown and Sevenson, 1988; Packer, 1975.) Similarly, given the popularity of Balanced Scorecards in the corporate sector and innate appreciation of performance multi-dimensionality in the public sector, it is likely that among those R&D that do adopt a Balanced Scorecard, corporate R&D labs are more likely to adopt a Balanced Scorecard approach while public-sector labs are more likely to adopt more generic multi-attribute performance frameworks. Academic labs would tend to trail these two sectors, as they have usually done with management fads in the past (Birnbuam, 2000). This leads us to the first set of hypotheses of our research:

  Hypotheses Set # 1:

  - HS.1.1—The use of Balanced Scorecard-type approaches is not prevalent among R&D organizations.
  - HS.1.2—To the extent that Balanced Scorecard-type approaches are found employed in R&D settings, they are more likely to be used in corporate R&D settings than in public sector with the academia lagging behind the two.
  - HS.1.3—Multi-dimensionality of performance is deeply rooted in measurement philosophies of R&D organizations.
  - HS.1.4—Multi-attribute performance measurement systems are more likely to be found in public sector than in private-sector labs with academic labs lagging behind the two.

Alongside the above hypotheses, we are also interested in developing a sense of how (and in what ways and why) are those organizations that do implement a Balanced Scorecard different from those that do not. For example: do R&D organizations implicitly "balance" across performance dimensions even though they may not explicitly do so?

- ***HS.2—Balance, Accessibility, Participation, and Transparency in Strategy-Making—*** One of the most distinct qualities of the Balanced Scorecard is the centrality of strategy in determining the performance measurement system (Kaplan and Norton, 2001.) It emphasizes a differentiated view of strategy best epitomized by Porter's (1996) assertion that "strategy is about making choices" and Treacy and Wiersema's (1995) three mutually exclusive customer value propositions. The literature on traditional Balanced Scorecards encourages the organizations to pick and choose a strategic posture instead of trying to be something for everyone (Kaplan and Norton, 2001.) Despite its importance, however, none of the literature on Balanced Scorecard (e.g. Kerssens-van Drongelen et al., 1999; Loch and Tapper, 2003; and Bremser and Barsky, 2004) explicitly talks about this differentiated view of strategy. We develop and test an adaptation of Kaplan and Norton's generic strategies to assess whether R&D organizations take a differentiated view of strategy.

While strategy itself is important, so is the strategy-making process. The Balanced Scorecard requires a process that is accessible, participative, and transparent, and that explicitly recognizes multi-dimensionality of organizational performance. This is done through explicit utilization of cause-and-effect performance models and strategy maps that operationalize the organizational strategy down to the level of individual employees (Kaplan and Norton, 2004.) Others have also emphasized the importance of the strategy-making processes and asserted that organizations with highly capable (multi-modal) strategy-making processes outperform those that have single-mode or less capable ones (Hart and Banbury, 1994). Wagner (1994) finds that participation [in strategy-making] can have a statistically significant difference on individuals' performance and satisfaction. Segev (1987) has emphasized the need for matching strategy-types with strategy-making process with the implication that a Balanced Scorecard-type approach to strategy-making must be accompanied by the appropriate process to exploit its full potential. This gives rise to a set of hypotheses:

Hypotheses Set #2:

- HS.2.1—R&D organizations, in general, and those implementing a Balanced Scorecard, in particular, take a differentiated view of organizational strategy.

- HS.2.2— R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, that implement an accessible, transparent, and participative strategy-making process report higher satisfaction and realize greater improvements in performance.
- HS.2.3—R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, incorporate the notion of performance multi-dimensionality in their strategy-making processes
- HS.2.4— Transparent and participative strategy-making processes in R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, provide opportunities for strategic ("double loop") learning for stakeholders.

- ***HS.3—"Balance" among Performance Dimensions in the Balanced Scorecard***—As discussed earlier, the acknowledgement of performance multi-dimensionality and the desire to achieve balance across these dimensions is the central tenet of the Balanced Scorecard (Kaplan and Norton, 1996.) This is also a recurring theme in the R&D management literature (e.g. Brown and Sevenson, 1988; Gretchen and Malone, 2002.) Kaplan and Norton (2001) and others, however, aspire for a balance of a particular kind i.e., one that explicitly acknowledges the existence of an underlying cause-and-effect model and uses it to communicate and drive performance itself. This also allows organizations to bring "laser-like" focus on the strategy and forces them to "measure the strategy". In the absence of any evidence to the contrary, we would assume that R&D organizations are not dissimilar to their corporate counterparts in this respect. This leads us a to a third set of hypotheses for this research:

  Hypotheses Set #3:

  - HS.3.1—R&D organizations implement performance measurement architectures that acknowledge performance multi-dimensionality and are balanced across various performance dimensions?
  - HS.3.2— R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, implement performance measurement systems that use explicit cause-and-effect performance models to "measure the strategy."
  - HS.3.3— R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, that incorporate the qualities described in HS.3.1 and HS.3.2 report higher satisfaction and realize greater performance improvements

- ***HS.4—Use of Incentives to Align Individual and Organizational Performance***— While the founders of the Balanced Scorecard movement do not provide explicit instructions to link incentives with the strategy-performance architecture, they do anticipate this to be a major challenge for the future development of the idea (Kaplan and Norton, 2002.) Others have also discussed the importance of rewards in implementing organizational strategy, especially, for top-management executives (e.g. Stonich, 1981.) More traditionally, though, top-managers constituted the only level of organizational hierarchy that enjoyed strategy transparency and visibility. The Balanced Scorecard provides an opportunity to make strategy visible to each and every employee within the organization through participation in, and operationalization of, organizational strategy.

There is considerable literature on incentives in R&D organizations (e.g. Pelz and Andrews, 1966; Kim and Oh, 2002.) Others have indirectly implied the importance of financial incentive for performance through returns to skill and performance (e.g. Gibbs, 2001; Diamond, 1986.) Despite these contributions, however, the debate is more or less divided between the proponents of intrinsic and extrinsic factors of motivation with the former preferring professional over monetary incentives rather than the other way round. We are interested in assessing whether R&D organizations believe in the ability of incentives to influence performance (i.e. do they think "incentives matter"?) and the mix of incentives employed at various types of R&D organizations. The following hypotheses motivated our investigation of the incentives issue in the context of Balanced Scorecard.

Hypotheses Set #4:

- HS.4.1— R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, believe in the ability of incentives to influence individual performance (i.e. "they do think that incentives matter?".)
- HS.4.2— Of the R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, receptiveness and prevalence to incentives is highest among private-sector labs and lowest among academia with public labs falling in between these two extremes.
- HS.4.3— Of the R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, higher number of incentives ("incentives menu") would affect performance positively, incentives intensity would have a inverted "U" curve relationship

with performance, and the effect on performance would increase as we move from financial to professional incentives.

- ***HS.5—Quality of Alignment between Strategy, Performance and Incentives Systems—*** The quality of alignment between strategy, performance, and incentives systems and the larger organizational strategic management system is one of the central structural features of the Balanced Scorecard. Multiple sources of strategic guidance and conflicting and confusing messages from various components of organizational management systems can easily nullify the benefits of individual mechanisms that are well-designed and optimal individually but do not align well with the rest of the strategic management system. Also, organizations that selectively adopt certain components of an overall strategy-performance-incentives system for implementation may not fully exploit the benefits of such an implementation (Hackman and Wageman, 1995.) The following hypotheses allow us to explore the effect of the quality of alignment between strategy-performance-incentives system with the perceived or realized benefits from the same:

  Hypotheses Set #5:

  - HS.5.1— Of the R&D organizations implementing Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, those with better-aligned strategy-performance-incentives systems report greater satisfaction and realize higher performance improvements.
  - HS.5.2—Of the R&D organizations implementing the Balanced Scorecards, in particular, and multi-attribute performance measurement systems, in general, those that selectively adopt certain components of the overall strategy-performance-incentives frameworks report lesser satisfaction and realize fewer performance improvements than those that do not do so.

This is an ambitious analytical agenda and being able to convincingly answer all of these questions is a tall order. Real life organizations are much messier than one would rather have them. While, several of the questions that we have identified above may be answered directly, others may only be addressed in an indirect way. For example, for some of these questions, one may be able to gather good enough data to isolate the phenomenon of interest while for others one may need to conjecture about possibilities or simply acknowledge lack of analytical capability.

We must develop an analytical approach that captures the various elements of this ambitious analytical agenda, thus allowing us the possibility of studying the phenomenon of interest in its multifarious and intricate details. Next, we outline such an approach.

4.2—THE RESEARCH METHODOLOGY

Given the highly contextual and multifarious nature of the research problem and the broad array of questions, namely, exploratory, descriptive, and explanatory, that interested us[61] we adopted a multi-pronged methodological approach. Using multiple analytic approaches was a natural choice given the strengths and weaknesses of the various approaches that could be deployed. For example, while a mail survey could bring generalizability to our results by allowing us to conduct statistical analysis on the constructs of interest, it had the obvious problems of survey response rates, limited ability to capture contextual details, and inability to convey difficult concepts in limited survey space and time. Case study analysis, on the other hand, scored better than the mail survey in its ability to incorporate rich contextual details in the analysis but was more costly conduct and analyze. Factors extraneous to the research problem per se (e.g. budgetary constraints, time constraints, and feasibility) also dictated the final choice of the analytic approach.

We used a multi-tiered research methodology whose various components, not necessarily in a sequential order, are presented below (please see Figure 4.1 for a graphical snapshot of the methodology):

4.2.1—The Literature Review(s)

The first part of the multi-pronged research methodology consisted of an in-depth literature review. Given the nature of the problem in question, this part of the methodology drew upon literature from a number of different topical areas, namely, the theory and practice of R&D and innovation management, organizational behavior and incentives design, performance management and measurement systems, quality assessment, and the measurement science in general. Lessons were also drawn from the practice of performance assessment methodologies at a selected group of existing R&D and other organizations (e.g. health, education, and governance.) The literature review was carried out as a set of four inter-related activities, namely, a preliminary literature review, a detailed literature review, an on-going literature review, and a reflective literature review. The following tasks were achieved as a result of the literature reviews at various stages of research:

[61] i.e. the how-many or the prevalence (e.g. "how many R&D organizations implement a Balanced Scorecard"), the how or the nature (e.g. "how do the Balanced Scorecards in R&D differ from those in the business organizations"), and the why or the effectiveness (e.g. "Do R&D organizations implementing the Balanced Scorecard perform better than those that do not? And why? etc.)

- The ***preliminary literature review (LitR1)*** was conducted before the start of the project. It helped develop the basic understanding of the organizational issues confronting R&D labs and understand the prior art in the area of R&D performance measurement and management thus putting the problem in a broader context. Using the basic insights gained during LitR1, we were able to formulate the policy and research questions for the proposed research.

- The ***detailed literature review (LitR2)*** began right after the first draft of research questions was developed. During the LitR2 phase, we developed a detailed picture of performance measurement in R&D. This entailed addressing questions like the definition of "success" used over time, the performance dimensions valued by practitioners and academics, various methodologies used over time, and shifting consensus in the literature over time.

  We also developed an appreciation of various R&D measurement approaches and paradigms (e.g. MBO, TQM, DFSS, and the Balanced Scorecard) that have been used in R&D labs over time with a view towards identifying their qualities, benefits, and limitations. We also developed an in-depth understanding of the Balanced Scorecard as theorized and practiced its founders as well as their followers and critics. This resulted in the formulation of the generic analytic framework that was used during later analytic phases. The detailed literature review contributed significantly to refinement of the research and policy questions, the development of the survey instrument and the case study protocol, and the theoretical hypotheses for the research.

- The ***on-going literature review (LitR3)*** helped inform important aspects of other methodological approaches and other problems as and when they arose during the implementation of the methodology. As we implemented our proposed methodology, we had to constantly refer back to the literature for guidance and advice. On-going literature review also helped us keep abreast of the changing state-of-the-art in the area of R&D performance measurement and incorporate newer findings in our analysis.

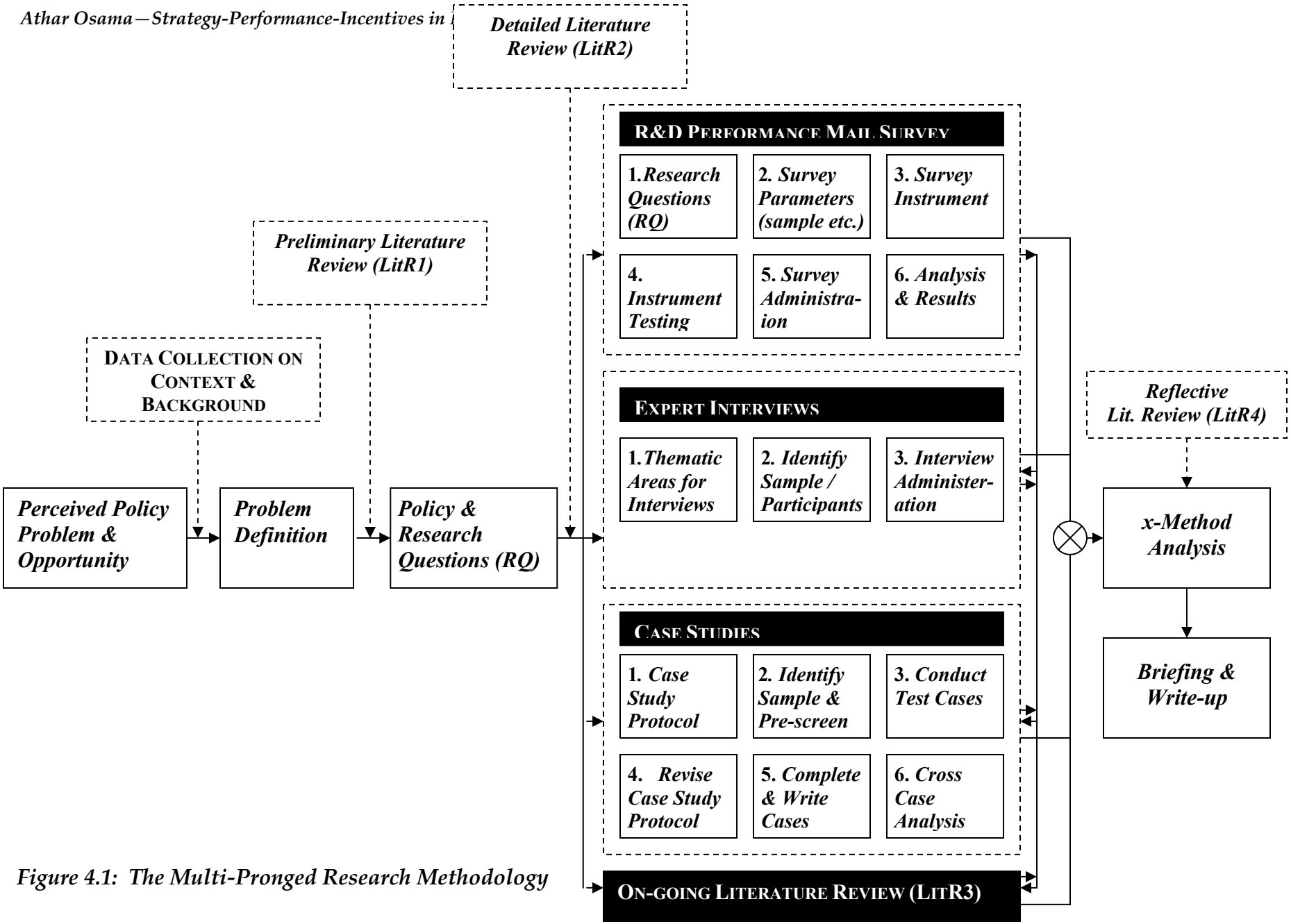


Figure 4.1: The Multi-Pronged Research Methodology

- Finally, a ***reflective literature review (LitR4)*** started after the culmination of the fieldwork and survey administration. It helped place our findings in the broader context of the R&D management and performance measurement literatures. We also used this opportunity to double-check and cross-check the findings of our own research against other research in the field and draw conclusions about the reasons for the same.

This multi-stage literature review process, with each stage having its own objectives and approach towards looking at the literature, provided us with opportunities to continually enrich the implementation of the research effort in line with developments in the literature. The reflective literature review, in the tradition of Kerssens-van Drongelen (2001,) facilitated theory building and allowed us to connect the findings with the broader literature.

4.2.2—A Mail Survey of R&D Organizations

The second phase of the research methodology comprised a mail survey of major R&D performers in the United States. The survey addressed issues like prevalence of various types of R&D performance assessment schemes, their uses and expectations, their effect on performance itself as well as the their relationship to strategy formulation and incentives in these organizations. It also checked for the use of various components of the traditional Balanced Scorecard in our population of interest.

4.2.2.1—The Rationale for the Mail Survey

Studying R&D organizations and scientists through survey methods has been one of the more difficult of the tasks confronted by those working in the realm of organizational studies (Pelz and Andrews, 1976.) The difficulty of getting senior R&D scientists and research managers to spare the time necessary to answer a survey of any meaningful size and scope is formidable. Low response rates and inability to cover a sample frame that is representative enough can lead to serious methodological challenges and issues. During the detailed survey of the literature, we encountered only a few studies of R&D organizations that had response rates upwards of 50% for individually administered surveys--a standard considered somewhat respectable in organizational studies (Aimen-Smith and Markham, 2005). Hence researchers have made an effort to use captive audiences to validate their hypotheses (e.g. Tipping et al., 1995; Wang et al., 1999.) Majority of the work in this area comprises practitioner interviews and case study research (e.g. Werner and Souder, 1997; Hurley, 1997.) There are some exceptions. Figure 4.2 presents a snapshot of some of these studies with response rates and key features of the survey methodology used.

FIGURE 4.2: EXAMPLES OF SURVEY RESEARCH IN THE POPULATION OF INTEREST			
Study Name/Authors	***Sample Frame***	***Response Rate***	***Survey Techniques***
UNIT OF ANALYSIS: R&D ORGANIZATION			
National Comparative R&D Project (1984-99) (Crow & Bozeman, 1998)	*(i) 825 Energy Sector Labs (ii) 1012 R&D Labs in U.S.*	*(i) 30% (ii) 72%*	*(i) Two mailings, 1 Telephone Reminder (ii) Telephone survey*
R&D Measurement Study (Jamsen et al., 2001)	*340 R&D Managers of Finnish Industrial Cos.*	*24.12%*	*Single iteration of a mail survey*
PM Systematic Design App., I.C.K van-Drongelen (2000)	*240 R&D Managers of Dutch Cos.*	*21%*	*Mail survey*
Pharma/Biomed CSFs Study Omta (1995)	*76 Biomed Research Units 20 Pharmaceutical Cos.*	*53% 70%*	*Prior approval of Entities' Participation was gained*
IRI's Biennial R&D Trends IRI (2002)	*227 Paying IRI Members*	*34%*[62]	*Mail Survey*
UNIT OF ANALYSIS: INDIVIDUAL SCIENTIST/ENGINEER			
Social Factors in Research Performance (1967-76) Pelz & Andrews (1976)	*1311 scientists and engineers in 5 locations*	*Not Reported*	*Convenience Sample Self-administered paper surveys*
Effects of Rewards Types Chen et. al. (1999)	*1169 R&D Professionals across 30 Companies*	*Not Reported*	*Surveys distributed to Lab Directors*
Size & Productivity Study Yeaple (1992)	*152 R&D professionals in 5 companies*	*57%*	*Mail survey, $1 Token Inducement*
Phd Trajectories Study Mangematin (1999)	*869 New PhDs from a French Institute*	*45%*	*Single iteration of a mail survey*
Phd Trajectories Study Coredero et al.,(1994)	*Employees of 24 Cos. 3163 returned surveys*	*55% (28-87%)*	*Single iteration of a mail survey*

Clearly the response rates of organization-level studies are much lower than what would be necessary to generalize to the larger population. Despite concerns for response rate and sampling bias etc., we decided to adopt the survey approach as one of the components of our research methodology for several reasons:

First, the mail survey would provide us with a sense of the universe—and to introduce us to the diversity within our target population. It could give us additional data points to look at and provide us with the ability to discover things, albeit with some degree of inaccuracy, that we would not have been able to do without a mail survey. For example, issues like the prevalence of Balanced-Scorecard-type measurement systems or the use of strategy as a differentiating tool

[62] Corresponding response rates were 48% in 2001, and 38% in 2000.

within R&D organizations could not be addressed, in a convincing manner, without actually surveying the landscape. Thus, a mail survey—irrespective of its biases—would be of considerable value in answering questions of the type described above.

Second, a mail survey would be valuable as a means to support or challenge the findings of other analytic approaches i.e. case studies or practitioner interviews. For example, if an overwhelming finding in the survey would support a common conclusion across several of our case studies, we might feel more comfortable making both a statistical and a theoretical generalization. On the contrary, if a regularity across our case studies failed to show up across organizations represented in our mail survey, we would be forced to be more humble and cautious about making a generalization[63].

Third, the mail survey could possibly provide us with an ability to make some correlational and causal inferences about a series of dependent (e.g. organizational performance) and independent (e.g. structure, strategy, performance measurement systems, incentives systems etc.) variables. These questions are highly relevant to the purpose of our analysis and the ability to answer them depends on a host of factors including, but not limited to, the response rates and sample sizes, potential biases, validity, and generalizability of various constructs etc.

Finally, the mail survey could allow us to prospect and pre-screen organizations that could potentially become subjects of detailed case studies. Keeping in view the above purpose, the survey instrument specifically asked the respondents if they would like to contribute to the study in a more substantive way by agreeing to a practitioner interview or become a subject of a full-fledged case study. About a quarter of our respondents agreed to one of the two possibilities thus enriching the sample frame from which we ultimately chose our case studies and allowing us the possibility to pre-screen and qualify the six organizations we chose to do detailed case studies on.

For all of the above reasons, we included a mail survey as one of the several methodological approaches for the proposed study. The following subsections describe the technical details of the survey design and instrument.

[63] This analytic scheme is particularly potent for extreme results/occurrences than it is for findings at the margin. For example, if we did not find even a single example of an R&D organization implementing a Balanced Scorecard among our six cases, we would be tempted to conclude that the Balanced Scorecard-type approaches are not R&D organizations. Here, the survey findings, if they do not conform to that conclusion, might help rectify a misperception.

4.2.2.2—The Population and Sample Frame

In keeping with the objectives of the mail survey, we developed a sample that, although non-representative from a technical standpoint, did provide a good coverage of the population of interest, namely, private sector labs, university research centers and labs, and public sector labs. There are more than 14,000 R&D Labs in the United States, ranging from small job shops and one-man operations to large multi-billion dollar federal labs (Crow & Bozeman, 1998.) While this is a truly diverse group in terms of type of work performed, organizational structure, mission, funding mechanisms etc., not more than 1000-1500 of these are significant in any measure of research inputs, outputs or outcomes. Our sample frame, therefore, comprised significantly large performers of R&D in public and private sectors and the academia. We used several data sources to identify such organizations, namely:

a) 220+ Members of Industrial Research Institute (IRI) that comprises the largest organized body of performers of industrial research in the US;
b) 250+ Research Extensive and Intensive Universities[64] in the US that comprises the largest and most significant research universities in the US;
c) 400+ Members of the Federal Laboratories Consortium that comprise a significant number of research performing organizations in the public sector in the US;
d) Fortune 1000 companies operating in technology and research intensive areas;
e) Fortune's 100 list of Fastest Growing Companies operating in the high-tech. areas;
f) 100+ companies cited by the Technology Review Magazine that comprises largest R&D performers (TR100 list);
g) Search results produced under the SIC categories "873xx" (Research and Development) using the Pro-USA Database of US Businesses; and
h) Search results produced by using the phrase "corporate lab" or "R&D lab" in Corporate Affiliations Database.

These sources combined produced about 1082 unique organizations that constituted the sample frame for the survey. The cover letter accompanying the survey instrument was addressed to "Chief Scientific, Technology or Research Officer" for the corporate labs, "Research Director" for the public sector labs, and "Provost or Vice President of Research" for the academia. These respondents were also encouraged to circulate the survey to other Research Directors or Managers within their organization or sister-organizations—thus encouraging multiple entries from the same organization. The mailing addresses for these organizations were collected from various sources, namely, business databases (e.g. Lexus Nexus, Pro USA etc.) for the corporate labs, Federal Laboratory Consortium's Members Directory for the public sector labs, and Carnegie Classification Lists and websites for the Universities.

[64] Research Intensive and Extensive Universities are defined by Carnegie Classification of Educational Institutions. Please see www.carnegiefoundation.org/Classification/ for details.

4.2.2.3—Survey Instrument Development and Administration

The development of the survey instrument began with an in-depth review of the relevant literature. We looked at other surveys conducted in the areas of R&D management and measurement to try to identify possible similarities and overlaps (e.g. Crow & Bozeman, 1998; I.C.K. van-Drongelen, 2001.) The final instrument was tested on an expert/representative of each of the diverse audiences (i.e. corporate R&D, public sector R&D, and Defense R&D communities) and the wording for some questions were altered and redundancies removed to improve the readability, consistency, structure, and content of the survey instrument.

The final survey instrument had six (6) distinct parts containing 29 questions (please see appendix B for the survey instrument). Figure 4.3 presents a graphical breakdown of the survey instrument.

FIGURE 4.3: R&D PERFORMANCE & INCENTIVES SURVEY INSTRUMENT

Section / Title	*# of Qs.*	*Types of Questions*	*Answer Format*
#0: Organizational and Personal Contact Information	*N.A.*	*Names, Position, Case study participation etc.*	*Blanks*
#1: Corporate Structure and Performance Information	*5*	*Ownership structure, employees, profits, expenditures, funding etc.*	*Multiple choice, Blanks*
#2: R&D Lab/Group/Team Structure and Performance Information	*9*	*Organization, expenditures, qualifications, performance etc.*	*Multiple Choice, Blanks*
#3: R&D or Technology Strategy	*3*	*Strategic posture, performance dimensions, degree of emphasis*	*Multiple Choice, Multi-choice*
#4: Performance Measurement Philosophy	*2*	*Potential and actual uses of performance information*	*Multiple Choice*
#5: Current Performance Measurement Systems/Approaches	*4*	*Assessment frameworks, duration, impact, metrics*	*Multiple Choice, Multi-choice*
#6: Performance Measurement Systems –Incentives Linkage	6	Compensation rules, individual and facility incentives etc.	Multiple Choice, Open-ended

Two versions (a longer, and a shorter) of the survey were developed and two mailings were conducted. The first mailing was carried out during June of 2003. As the response from the first mailing tapered off, a second mailing followed around mid-July. Each potential respondent was mailed a copy of the survey containing seven (7) pages of questions, a detachable page of instructions, and a page-long data dictionary containing explanations of terms used in the survey. An email alias was established to respond to any queries that the respondents might have while answering the questionnaire. The names of the principal investigator, the study director, and three study advisors were clearly indicated on the cover letter along with the telephone numbers, fax numbers, and email addresses of the principal investigator.

The respondents were encouraged to use any transmission means they deemed convenient for returning the questionnaire, including fax and surface mail. However, a postage-paid envelope was included in the packet to indicate the preferred method. The respondents were offered an

advance copy of the study results as an incentive to participate in the survey. Participation was completely voluntary and anonymous. Respondents were, however, required to disclose their identity if they volunteered to become a subject of a case study or practitioner interview. Prior to the mailings, the survey instrument and the accompanying cover letter went through a review by the Human Subjects Protection Committee (HSPC) at RAND.

4.2.3—R&D Practitioner/Manager Interviews

The third component of the overall research strategy comprised interviews with practitioners of R&D measurement and management (e.g. Vice Presidents or Chief Scientific, Technology or Research Officers or Directors of Research of R&D performing companies, Directors of Research of Public Labs, Vice Presidents of Research at major research universities, research managers and senior scientists) as well as experts on the subject (e.g. academics and consultants interested in research management and measurement.) We used following methods to identify R&D practitioners and experts:

- ***Convenience sample***: We asked known experts in the area to identify others active in the field, attended conferences to make contacts with practitioners, and gleaned through past publications in professional journals to identify potential targets for interviews and contacted them through an email or telephone.
- ***Opt-in to the Study:*** Several of the survey respondents also opted to participate in a detailed interview. We also used a pre-screening interview for the case study process to solicit expert and practitioner views on the subject of interest.

Each interview comprised 45-60 mins. and focused on six key areas, namely, general views about performance measurement and its uses, nature of the strategy formulation process, specific notions of performance measurement being used at the interviewees' organization, use of incentives to align individual motives with organizational goals, problems faced during implementation of new performance measurement systems and ways to get around those problems, and interviewees' perceptions of the appropriateness of the balanced scorecard in R&D.

The expert-practitioners' interviews could contribute to the research in several ways, namely, they could provide rich contextual details of the problem not possible in a mail survey, the ability to probe deeper into qualitative issues that became important during the course of the study. an opportunity to seek clarifications and explanations of unexplained findings to refine our own ideas and hypotheses about the subject matter, and developing a rough sense of the extent to which our survey respondents or case study subjects were representative of the larger population.

4.2.4—Detailed Case Studies of R&D Organizations

The case study analysis was the mainstay of our proposed research design. Case studies stood out as preferred research methodology for several reasons, many of which are highlighted in Yin (1987, 1993), for example:

a) Our research questions were aimed at exploring the "hows" and "whys" of a real world complex phenomena;
b) It was difficult to isolate the phenomenon of interest (i.e. performance and incentives in R&D organizations) from its context (i.e. type of organization, nature of R&D work, size, culture, objectives and institutional form etc.)
c) Due to the contextual complexity, the variables of interest were much more numerous than the data points that can be feasibly collected thus making regression-like methodologies difficult-to-implement; and
d) It was both possible and necessary to use multiple sources of data (e.g. documents, artifacts, interviews, and observations.) and bring them to bear on the research.

The case study process started with a theory building exercise that identified a set of descriptive, exploratory, or explanatory constructs and hypotheses for the overall case study analysis. The study adopted a multiple case design to allow us to learn from and across several organizations. This was important primarily because we thought that the main subject of this investigation, namely, a Balanced Scorecard-type multi-attribute performance measurement system, may not exist in an ideal form in any one R&D organization that we knew of but components of such frameworks do exist in a number of organizations.

A multiple case study design could allow us to piece together fragmented evidence from a number of performance measurement systems found in practice and make a meaningful judgment about the appropriateness of putting together a more complete system. Performance measurement system within an individual organization was the main unit of analysis and the focus of our investigation. However, we believed that a performance measurement system cannot be isolated from its overall context (i.e. social and behavioral factors, for example, environment, motivations, competitive landscape,) on the one hand, but more importantly strategy and incentives systems, on the other. This conjecture was further strengthened by the central importance strategy occupies in the Balanced Scorecard literature (Kaplan and Norton, 2001.)Therefore, the study, while focusing on the performance measurement systems, also investigated its linkage upstream (with strategy) and downstream (with incentives) as well as with the overall organizational context.

4.2.4.1—The Study Questions for Case Study Analysis

FIGURE 4.4 (A): STUDY QUESTIONS—INDIVIDUAL VS. CROSS CASE ANALYSIS	
Individual Case Studies	***Cross Case Analysis***
How do organizational R&D leaders think about performance and incentives in R&D?	*How do R&D leaders and managers, in general, think about strategy-performance-incentives?*
How well-developed, rigorous, participatory, and "balanced" is the strategy-formulation process at the organization?	*Do organizations implementing "balanced" and participative strategy-planning systems tend to perform better than those that do not?*
Does the performance measurement system reflect balance across performance dimensions? Does it follow from strategy?	*Do organizations that use performance measurement systems that are "balanced" and aligned with strategy do better?*
What type of incentive systems exist, and how do they align individual motivations with organizational goals?	*Do organizations that clearly link performance with incentives/rewards tend to do better than those that do not?*
Do strategy-performance-incentives systems improve organizational performance?	*In what ways do institutional environments (e.g. type of research, organization) affect strategy-performance-incentives systems?*

FIGURE 4.4 (B): DATA COLLECTION REQUIREMENTS & INTERVIEW SCHEDULE	
Section of the Case Study	***Suggested Interview Contact(S)***
Corporate (or Lab's) Technology/R&D Strategy *(deals with organizational strategy) – 1 hour*	VP of Technology, Chief Scientist/Engineer OR Director of R&D
R&D Performance Measurement System *(deals with dimensions, system design, metrics) —1 hr.*	Mid. Mgmt.-level individual dealing with performance measurement system in R&D
Performance-Incentives Linkage *(deals with incentives-rewards & performance)—1 hr.*	Mid Mgmt.-level individual (in HR perhaps) dealing with incentive design and compensation system issues
R&D Environment, Motivation & Incentive Effect *(deals with assessing S&E motivation factors)—1 hr.*	Mid level Bench Scientist(s) or Engineer(s) working for a few years preferably during the introduction of PMS
General Info. & Coordination *Appox. 4 hrs. over the case study duration (~0.5-1 mo.)*	A junior level individual who might be able to help locate information, and answer relevant questions.

The case study evidence could be utilized at two levels. The first of these levels is that of an individual organization. Every organization that we studied was unique—not only it terms of its goals, organization, and management style but also in the strategy-performance-incentives systems it utilized and how it utilized them. At the level of an individual case, we sought to answer questions relating to the strategy-formulation, performance measurement, and incentives systems within each of the organization studied. The key contribution of this case-level analysis could be to allow us to study these constructs of interest within the highly individualized and unique

circumstances of that particular organization. The second of these levels of analysis is the cross case comparison where we tried to identify the similarities and differences across several organizations and draw generalized conclusions about the constructs of interest. Figure 4.4(A) presents a summary of questions we sought to address at each of these levels.

In accordance with the questions identified above, each case study focused on five key areas:

a) ***General perceptions of performance measurement in R&D***
b) ***The strategy-formulation process within the organization***
c) ***R&D performance measurement at the organization***
d) ***Incentives alignment within the R&D organization***
e) ***Implementation dynamics and issues***

To cover these five focus areas, we sought several telephone or in-person interviews within each case study organization. These are summarized in Figure 4.4 (B).

4.2.4.2—The Study Design and Plan for Analysis

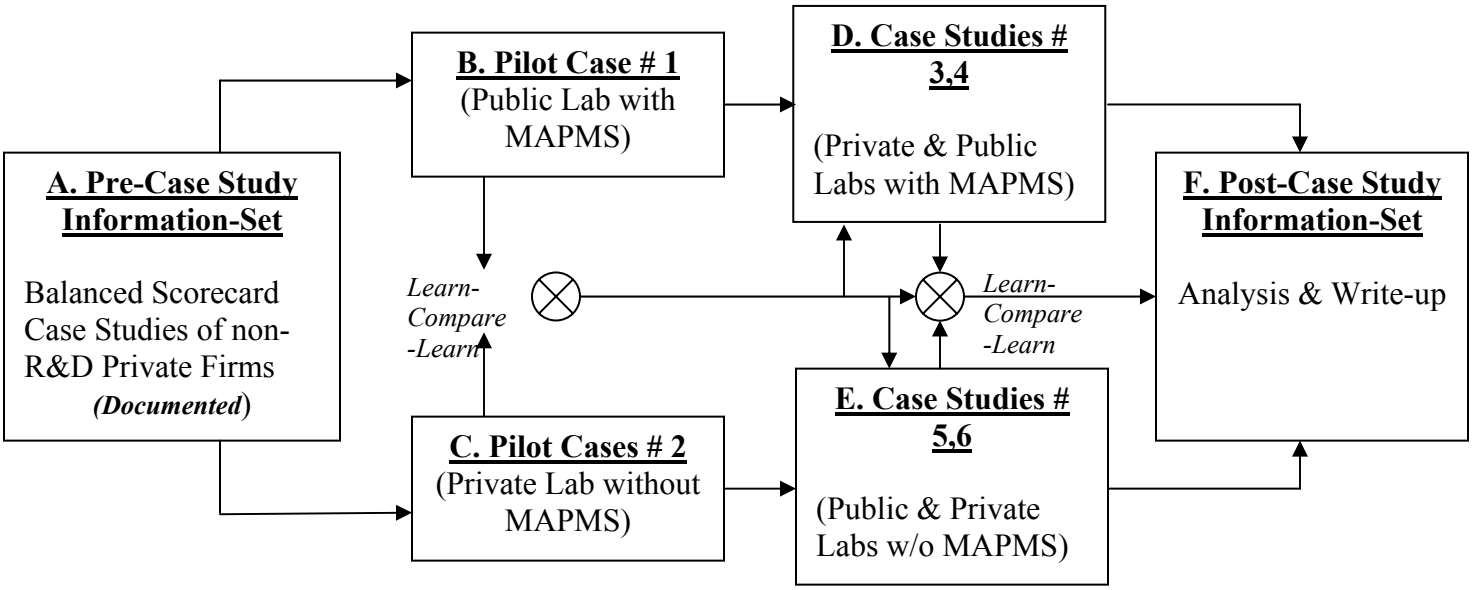
In addition to developing a broad understanding of the constructs identified in the questions (Figure 4.4-A), the central concern of our investigation was the Balanced Scorecard-type multi-attribute performance measurement systems. We started our research with the basic premise that while the Balanced Scorecard itself has been applied with some degree of success to business world—its application has been fairly limited in the R&D environment[65]. This presented a challenge for us in designing a study that could provide us an opportunity to make conclusions and recommendations about a phenomenon that has not yet taken roots among our population of interest.

As the study progressed, however, we discovered that we were not in a totally uncharted territory. We learnt that a number of R&D organizations have —for quite a long time now—experimented with the notions of multi-attribute performance and scorecards—separately, if not as a part of an integrated performance measurement system, like the Balanced Scorecard. Many, in fact, most of the R&D organizations that we came across have used performance measurement systems that are definitely somewhere between the two extremes (i.e. uni-attribute and multi-attribute.) We also found several examples of Balanced Scorecard-type performance measurement systems in place and decided to exploit this variation to draw some conclusions.

The Figure 4.5 presents the case study design in a graphical manner. The key strength of the proposed design was its ability to use existing knowledge denoted by information set-A (i.e. ,Balanced Scorecard case studies in general business contexts) in conjunction with evidence gained

[65] Telephone interview with Alan Fusfeld (Sept, 2003) and conversation with Drs. Robert Kaplan and David Norton (December 2002).

from case study analysis of R&D organizations that used multi-attribute performance measurement systems (denoted by information set-B and D) and otherwise (denoted by information-set-C and E) under different organizational contexts (e.g. public and private sector organizations) to systematically compare the similarities and differences between these groups of organizations.



COMPARISONS	KEY ANALYTICAL OBJECTIVES ACHIEVED
A-B	Sheds light on differences between MAPMS in business vs. R&D
A-C	Sheds light on why R&D might be different from normal business firms?
C-B	Sheds light on how R&D firms using MAPMS are different from those that don't?
B-D	Builds upon findings of the pilot in B. (achieves literal replication)
C-E	Builds upon findings of the pilot in C. (achieves theoretical replication)
D-E	Builds upon findings of C-B
D-E-F	Tests hypothesis at the cross case level.

FIGURE 4.5: THE PROPOSED CASE STUDY DESIGN & ANALYTIC PLAN

The analytic plan was grounded in the principles of literal and theoretical replication within and between cases. The concepts of literal and theoretical replication—drawn from the work of Robert Yin (Yin, 1987, 1993)—hinge on the idea of validation of the underlying theory through a

FIGURE 4.6—THEORETICAL MODEL/PRIORS FOR CASE STUDY ANALYSIS	
RESEARCH ISSUES	**HYPOTHESES & THEORETICAL PRIORS[66]**
Acceptability and Prevalence of Balanced Scorecard	***HS1.1:*** Balanced Scorecard (BSC) not prevalent among R&D organizations ***HS1.2:*** BSC Use: Private→Public→Academic ↓ (usage goes down); ***HS1.3:*** MAPMS prevalent among R&D organizations ***HS1.4:*** MAPMS Use: Public→Academic →Private ↓ (usage goes down);
Balance, Participation, Accessibility, and Transparency of Strategy-Making	***HS.2.1:*** Organizations using BSC/MAPMS take differential view of strategy ***HS2.2:*** Organizations using BSC/MAPMS with balanced/participative strategy-making systems report higher satisfaction and realize greater performance improvements. ***HS2.3:*** Organizations using BSC/MAPMS incorporate performance multi-dimensionality in strategy-making ***HS2.4:*** Organizations using transparent-participative strategy-making engage in strategic ("double-loop") learning
Balanced Scorecard and Multi-attribute Performance in R&D	***HS3.1:*** R&D organizations implement measurement systems that are "balanced" across dimensions of performance ***HS3.2:*** Organizations using BSC/MAPMS use cause-and-effect modeling to "measure the strategy". ***HS3.3:*** Organizations using BSC/MAPMS that incorporate qualities of HS3.1/HS3.2 report higher satisfaction and realize greater performance improvements.
Use of Incentives to Align Individual with organizational performance	***HS4.1:*** Organizations using BSC/MAPMS believe in ability of incentives to influence individual performance ***HS4.2:*** Prevalence of Incentives: Private→Academic → Public ↓ ***HS4.3:*** Perceived or Realized Effect of Incentives: # ("menu")of Incentives ↑, Intensity of Incentives ↑↓, Financial→ Professional ↑
Alignment in Strategy-Performance-Incentives Systems	***HS5.1:*** Organizations using BSC/MAPMS with better-aligned strategy-performance-incentives systems report higher satisfaction and realize greater performance improvements. ***HS5.2:*** Organizations using BSC/MAPMS that "cherry pick" components of the overall strategy-performance-incentives system for implementation report lower satisfaction and realize lower performance improvements.

[66] The notation for this column is as follows: "Private→Public ↑" means that as we move from private sector to public sector R&D organizations, the acceptability and prevalence of performance measurement decreases.

FIGURE 4.7 – KEY METHODOLOGICAL CONCERNS REGARDING CASE STUDY ANALYSIS		
TEST OF QUALITY	**CASE STUDY TACTIC[67]**	**SPECIFIC APPLICATION (HOW USED?)**
Construct Validity	*Using multiple sources of evidence*	Multiple sources of evidence, described above, allowed us to not only check for inconsistency of definitions in key constructs but also improve internal validity through triangulation of evidence and corroboration.
	Establishing a chain of evidence	The study carefully documented the entire chain of evidence, and cross-referenced various methodological elements and facts to allow replication of findings and evidence in either direction. This minimized bias and improved construct validity.
	Having key informants review the case study draft report	Each draft case study report was shared with key individuals within the concerned organization to ensure integrity of data (not necessarily conclusions.)
Internal Validity	*Use of pattern matching*	For each case participant, based on an assessment of evidence prior to field study, a predicted pattern was defined and compared to one observed during the field study. A match in predicted and observed pattern strengthened the internal validity. Conflict in the above was used to update theory.
	Use of explanation building	The investigator sought to provide theoretical explanations in support of various observations made during the research and ruled out rival explanations (supported by evidence) to strengthen the internal validity of the study.
	Time-series analysis	We looked for opportunities to use quantitative data on performance measurement, performance, and incentives in a time-series fashion, where possible to support a causal linkage between the constructs of interest.
External Validity	*Use of replication logic in a multiple case study framework*	The study used replication logic (literal and theoretical) to demonstrate that results are generalization to a theory (not a population).
Reliability	*Use of case study protocol*	A detailed case study protocol was developed to allow documentation and standardization of procedures that can be used for replication of study results.
	Developing a case study database	A case study database containing objective evidence was developed to allow those interested to go beyond the researcher's conclusions and reconstruct the objective evidence on each case from which similar or different conclusions might be drawn by another researcher.

[67] We heavily draw upon Yin (1984) for the construction of this table.

theoretical replication within a single (or limited number of) study subject(s) rather than a statistical replication of results across a large number of subjects, as is done in survey research. Prior to starting the case study process, we developed an entire set of theoretical priors that we expected to be present among our case study subjects. Figure 4.6 provides a brief representation of these theoretical priors. It also highlights the type of research (i.e. exploratory, descriptive, or explanatory) that was pursued within each of the six areas.

The case study part of the research comprised two parallel analytical streams. At the level of an individual case, each participating organizations was be analyzed using methods like *pattern matching* (i.e. comparing the observed pattern of constructs with a pre-defined set of pattern predicted by the theoretical model of how things should look like,) *explanation building* (i.e. explaining the possible reasons for observed deviations,) *time-series analysis* (i.e. using longitudinal data on performance to compare pre-post-intervention results, when possible,) and *analysis of embedded units* (i.e. in the context of the current study, to compare differential implementation of performance assessment schemes under differing sub-organizational context but under the same context of the larger organization). A separate report having a descriptive and an analytic section was developed for each individual case subject. The cross-case analysis attempted to consolidate the evidence found across multiple case studies into a coherent view of performance measurement in R&D.

Ensuring construct validity, reliability and generalizability of results in the case study approach is a major challenge. However, one can take adequate precautions to ensure that the case study method comes up to the expectation of rigorous defendable research (Yin, 1993). The steps taken to improve construct, internal, and external validity and reliability are highlighted in Figure 4.7 (above).

4.2.4.3—Implementation of Case Study Analysis

A total of six case studies were conducted. Case participants were carefully selected on the basis of a couple of criteria, one pragmatic and the other methodological. The pragmatic criteria was used as a first screen to identify organizations that were available and willing to provide a required level of access and time-commitment for the purpose of becoming a case study participant. The methodological criteria was used as a second and final screen to further narrow down the list of potential case study subjects to only those organizations that fitted, to the extent possible, the description or profile of subjects needed to implement the analytical strategy (see Figure 4.5.) A detailed case study protocol was developed to guide the data collection process. Two initial cases—one each from public and private sector organizations—were conducted as pilots and case study protocol and other materials were updated as a result of lessons learnt from these two cases. Each case study comprised a detailed examination of several sources of data which included

organizational documents, performance information, interviews with key individuals, published sources, web-accounts, annual reports, as well as site-visits, wherever possible.

The case studies were conducted during the August-October 2003 timeframe. Each case comprised an engagement of about 7-10 days—including al least:

- 1-day—pre-case study preparation (website, media reports, SEC filings, other material etc.),
- 1-2-days— on-site (interviews, site visit, data gathering etc.),
- 1-day—post-case study coordination (incomplete data, follow-up interviews, gaps in understanding etc.), and
- 5-days---for write up and analysis.

All cases were be done by a single investigator thus eliminating the need for training of field workers..

Collectively, these four methodological components allowed us to capitalize upon the strengths and weaknesses of each of these methodologies and provided us with a fairly detailed and robust picture of performance measurement in R&D in general, and multi-attribute performance measurement systems in particular. We present the key findings of our investigation in the three chapters that follow. In chapter-5, we present the findings of the mail survey. In chapter-6, we present individual case descriptions that allow us to go deeper into the highly contextual "stories" of each organization and analyze their strategy-performance-incentives systems in somewhat detail. We also discuss the impact on the actual performance of individuals that these systems might have had. In chapter-7, we present a cross-case analysis, combined with the generalizations from practitioner-expert interviews.

CHAPTER—5

DEVELOPING A SENSE OF THE UNIVERSE: LESSONS FROM A SURVEY OF R&D PERFORMERS

The mail survey of major public, private, and academic R&D labs in the United States was conducted between June-August, 2003 as per procedures described in chapter-4. Two consecutive mailings, approximately 5 weeks apart, were done to 1082 organizations. 55 surveys were returned unattended as a result of the first mailing[68], thus effectively reducing our sample frame to 1027 unique addresses. After the two mailings, 89 survey responses were received—87 of which were deemed fit for analysis. This amounted to an overall response rate of 8.47%. The responses represented a cross section of various organizational types (e.g. public, private, and university labs.) The Table 5.1 presents some key statistics on the survey respondents.

TABLE 5.1: STATISTICAL SNAPSHOT OF RESPONDING ORGANIZATIONS

Key Survey Statistics:	
Initial Size of Sample Frame	1082
Effective Size of Sample Frame	1027
Total # of Survey Responses Received	89
Total # of Survey Responses Deemed Unusable	2
Total # of Usable Survey Responses	87
Response Rate	8.47%
Breakdown by Organization Type (N=87):	
Private Sector:	30 (34.5%)
Public Sector:	26 (29.9%)
Academia:	31 (35.6%)
Breakdown by Type of R&D Work (N=84):	
Research Lab (>75% of work in Basic/Applied Research)	49 (59%)
Development Lab (<25% of work in Basic/Applied Research)	35 (41%)
Breakdown by Unit of Analysis or Type of Facility (N=78):	
Overall Company:	48 (61.5%)
Corporate R&D Lab:	11 (14.1%)
Divisional/Department Lab:	19 (24.35%)

[68] Either because the organization had ceased to exist, or the address was faulty, or the recipient organization had a policy of not accepting mail that did not address a specific individual by name.

5.1—ANALYSIS OF NON-RESPONSE AND NON-RESPONSE BIAS

Given the nature of our objectives for the survey component of our methodology, non-response bias does not constitute a very significant threat to subsequent analysis. We would, however, attempt to assess the nature and direction of non-response bias in our sample. Our response rates might have been affected by a number of factors—not the least important of which may have been the perception that the survey was specifically designed for private sector R&D organizations. We received several queries, especially from academic labs, asking us whether the survey was actually designed for them[69]. We answered the questions to the best of our ability, attempting to encourage the broadest possible audience to fill out the questionnaire. Table 5.2 presents some statistics on the within group response rates.

TABLE 5.2: WITHIN-GROUP RESPONSE RATES— BY ORGANIZATION TYPE

Total # of Organizations in the Sample Frame by Type of Organization	
Private Sector / Corporate:	522
Public Sector:	263
Academic/University:	257
Total # of Responses & Response Rates from Each Organization Type:	
Private Sector / Corporate:	30 (5.7%)
Public Sector:	26 (9.8%)
Academic / University:	31 (12.06%)

On the whole, we believe that the differences in within-group response rates may be heavily affected by a number of factors including, but not limited to, the differing motivations of those responding (e.g. public sector respondents might have been encouraged by the fact that the study was clearly aimed at helping public sector research managers, and university research managers might have been generally more sympathetic to a research endeavor and hence more likely to respond,) the workload of and the amount of time at the disposal of the responding individual, and the quality of contact information that we had rather than the perception of a bias in the design o the instrument itself. Omta (1995) discusses some of the reasons for non-response in our relevant study population and identifies lack of time or lack of interest (because of questionnaire weariness) as the two most important ones.

[69] Certain questions on the survey e.g. profit/revenue, sales from new products etc. might created an impression that the survey was designed for private sector R&D organizations.

One can get a rough estimate of the non-response bias by looking at the differences between early and late respondents. Late respondents are likely to be more similar in profile to the non-respondents (Omta, 1995.) Our survey design allows an opportunity to undertake that kind of an analysis. Specifically, we can compare the characteristics of early and late respondents by looking at those that replied before and after the second mailing of the questionnaire. Table 5.3 presents the characteristics of the early and late respondents in our sample.

TABLE 5.3— OBSERVABLE CHARACTERISTICS OF EARLY & LATE RESPONDERS

Organizational Attribute (Scale/Unit)	Entire Sample (N=87)	Early Responders (N=35)	Late Responders (N=42)
Ownership Structure (Private: Public: Academic)	30:26:31	31:20:48	34:28:38
Total R&D Expenditure (Continuous $US)	304 M	120M (26M)	433M (50M)[70]
Average Self-Described Performance Rating (1-4)	2.51	2.52	2.50
Breakdown of R&D Expenditure (Central: BU: Ext.)	29:50:21	27:55:17	34:37:35
Breakdown of R&D Exp. (Basic: Applied: Other)	30:28:37	34:28:32	28:29:40
Technical & Managerial Employment (1-5)	3.09	3.03	3.14
% of PhDs Employed	40%	47**	35
Salary (Scale 0-3)	1.74	1.79	1.70
# of Years of PM System in Operation	6.01	5.95	6.05
Perceived Impact of PM System	3.98	3.75**	4.125

The results indicate certain interesting differences between early and late responders. For example, early responders tend to be smaller than the late responders in terms of R&D expenditure—a difference that remained after accounting for outliers in the sample. Median annual R&D expenditure for early responders was half as much as late responders. Academic labs were over-represented and public sector labs were under-represented among early responders. There was no difference, on average, in self-described performance ratings of early responders and late responders. Early responders were significantly more likely to hire PhDs—a fact probably attributable to over-representation of university labs in this group. We did not find a statistically significant difference among the two categories on salary level or numbers of years the performance measurement system had been in place. The early respondents, however, tend to have a more negative perception of the impact of putting this system in place. The finding was significant at 10% level.

[70] Medians in the brackets to guard against outliers. The differences for medians is not statistically significant.

Another way to look at the relative propensities to respond to the questionnaire is to compare the organizational characteristics of those who replied to the first and the second versions of the questionnaire. The data is presented in Table 5.4. The columns titled "Intrustment-1" and "Instrument-2" refer to organizations that respondent to the first (longer) and second (shorter) versions of the questionnaire respectively. The column titled "Procrastinators" describes those organizations that responded to the Instrument-1 well after the second mailing (Instrument-2) has been done. These might be, as the name suggests, are those organizations that procrastinated.

TABLE 5.4— OBSERVABLE CHARACTERISTICS OF VARIOUS RESPONDENT TYPES

Organizational Attribute (Scale, Sample Size)	Instrument #1 (N=49)	Instrument #2 (N=36)	Procrastinators (N=9)
Ownership Structure (Private: Public: Academic)	30:25:49	36:39:25	44:33:22
Total R&D Expenditure (Continuous $US)	109M (34M)	567M (50M)[71]	89.5M (64.5)
Average Self-Described Performance Rating (1-4)	2.39	2.66	1.88**
Breakdown of R&D Expenditure (Central: BU: Ext.)	28:50:24	37:47:15	NA
Breakdown of R&D Exp. (Basic: Applied: Other)	33:30:33	27:27:43	19:26:49
Technical & Managerial Employment (1-5)	3.14	3.02	3.22
% of PhDs Employed	42	38	18.5
Salary (Scale 0-3)	1.70	1.79	1.66
# of Years of PM System in Operation	5.36	6.71	4.16
Perceived Impact of PM System	3.82	4.16**	4

There are few significant differences between the two groups, except that procrastinators rate their own performance considerably lower, somewhere between above-average and average. The difference is statistically significant at 10% level. As before, early responders (by this measure, as well) tended to over represent academic labs and under-represent public sector labs. They also tended to have smaller R&D annual expenditures. As before, they tended to have a more negative perception of the effectiveness impact of their performance measurement systems as compared to those who responded later.

Assuming Omta's (1995) assertion to be true, i.e. late responders are more likely to resemble non-respondents, we can speculate the properties of non-respondents and the potential for non-response bias in our sample. Specifically, non-respondents are more likely to be public sector labs, equally likely to be corporate labs, and much less likely to be academic labs than respondents. This also partially confirms our assertions derived from aggregate analysis of with-in sector response rates. Non-respondents and procrastinators are likely to be larger—perhaps more bureaucratic—

[71] Medians in the brackets to guard against outliers. The differences for medians is not statistically significant.

organizations. This can be partially explained by the additional difficulties in finding data to fill out the questionnaire in a large multi-divisional bureaucratic setting. Finally, non-respondents are more likely to be happier with their performance measurement systems. This also confirms Omta's findings on non-response bias in this population in that many respondents tend to believed their "research unit was already so successful that they did not see the need to enhance it by participating in [that] study" (Omta, 1995, p. 145.).

On several of the returned survey responses we also faced the issue of item non-response. This was expected[72], given the broad nature of the survey instrument and our attempt to appeal to the widest possible audience. It was thus quite logical for many respondents to find some of the items not applicable to their type of organization or R&D work. We considered the possibility of using statistical techniques often used to generate missing data (e.g. hot decking, replacing missing value with average values) but decided against the proposition due to sample size constraints and the diversity of the respondents.

This chapter is structured as follows. Section 5.2 begins by discussing the basic results of the survey. Section 5.3 builds upon this by presenting an analysis of the survey results. In this section we would specifically address two key issues of concern to this investigation, namely, the public-private-academic differences (5.3.1) and the differences in strategy-performance-incentives systems between high-performing and low-performing R&D organizations (5.3.2). In the next section (5.4), we discuss the key findings from the survey that are of material consequence to the central thesis of this research, namely, the appropriateness of a Balanced Scorecard-type performance measurement philosophy for the R&D settings.

5.2—THE SURVEY RESULTS

We analyzed the data in aggregates as well as in certain sub-categories of interest. The three main areas covered in the survey instrument, apart from organizational structure and performance variables, dealt with the three major constructs of importance for the study, namely, strategy formulation, performance measurement, and incentives systems within R&D organizations. At the aggregate level, for example, we were interested in finding out whether our subjects tried to differentiate themselves in terms of strategy or whether they tried to be "something for everybody". We were also interested in the types of measurement frameworks being employed at R&D organizations, how long have they been in place, and whether or not, in the opinion of the respondents, they had any impact on performance itself? Finally, we were interested in getting a feel

[72] Respondents were encouraged to respond to as many questions as they could and deemed relevant to their organizations.

of the types of incentives available to individuals working in these organizations, the rules for determining compensation, and the balance between monetary (financial) and non-monetary (professional) rewards etc.

In addition to the aggregate findings, we were interested in whether an organization's strategy formulation, performance measurement, and incentives systems were influenced by different organizational structures (e.g. public vs. private vs. academic labs,) or type of research carried out (e.g. basic and applied vs. development,) or self-described competitiveness of the organization (e.g. a world class lab vs. an average or a below average one) and how might these affect the performance of the organization itself. Before we present our findings, we would present a statistical snapshot of the organizations we studied.

5.2.1—Organizational Structure of R&D Organizations Studied

Our respondents represented an even cross section of organization types, institutional structures, and type of research work performed etc. Table 5.5 provides an anonymous snapshot.

TABLE 5.5: A CROSS-SECTION OF SURVEY RESPONDENTS (ANONYMOUS)

Private/Corporate Labs
- A mid-sized Specialty Chemicals Co.
- A major multinational Petrochemicals Co.
- A jointly-owned Lab in IT/Materials etc.
- A mid-sized specialty materials Co.
- A mid-sized foreign automotive Co. lab
- A major Telecommunications Co. Lab
- A multi-site lab of a major Aerospace Co.
- Corporate R&D Lab of a Microprocessor Co.
- A major medical products Co.
- Corp. R&D Lab of a foreign Electronics Co.

University-Academic Labs
- Several State Universities
- Several Private Universities

Public Sector (non-Defense) Labs
- USDA Agriculture Research Centers
- A National Institutes of Health Lab
- A US Dept. of Education Research Center
- NASA Research Centers
- Several DOE-managed US National Labs
- A Federal Aviation Administration Lab

Defense-related Labs
- Several US Army Research Centers
- Several Airforce Research Centers
- A Dept. of the Navy Research Center

In aggregate, these 87 organizations had combined annual R&D expenditures of $24.34 billion—quite a substantial sum when compared to the overall private and public research and development expenditure in the US. This amounted to an average R&D expenditure of around $297 million. There was, however, considerable variation in this figure with annual R&D expenditure figures ranging from $90,000 for the smallest organization in the sample to $1.5 billion for a mega lab

system[73]. For the 43 respondents that reported a figure for capital expenditure on facilities and equipment over a three-year period, the average came out to be around $32.5 million, again with a wide variation among them.

TABLE 5.6— SELECTED AGGREGATE ORGANIZATIONAL STRUCTURE ATTRIBUTES

Organizational Attribute (Scale, Sample Size)	**Mean (Std) / (Range)**
Attributes of the Corporate Parent:	
Total Employment *(Scale 1-5, N=49)*	3.16 (.92) / (1-5)
Total R&D Expenditure *(Continuous, N=43)*	$209 million (330M)/(200,000-1.3B)
% of New Products introduced in last 5 years *(1-3, N=25)*	1.72 (.89) / (1-3)
Break Down of R&D Expenditure	
Central Lab *(N=42)*	33.92% (38.09)/(1-100)
Business Unit Lab *(N=42)*	48.83% (43.26)/(1-100)
Contract or External R&D *(N=43)*	22.09% (33.74)/(1-100)
Attributes of the R&D Facility (Unit of Analysis)	
Total Employment *(Scale 1-5, N=85)*	3.09 (1.331)/ (1-5)
Total R&D Expenditure *(Continuous, N=82)*	$297 million (166M)/($90,000-$15B)
CapEx in Facilities & Eqpt. in last 3 yrs *(Continuous, N=43)*	$32.5 million (878M)/(0-$500M)
Average Salary of Technical Professionals *(0-3, N=84)*	1.77 (.75) / (1-3)

In terms of the overall size of employment in the parent organizations, of the 49 organizations that reported the statistic, the respondents approximately had a bell-shaped curve with about 77.5% of the organizations falling within the middle two categories on a 5-point non-linear scale[74]. For employment within the R&D facilities surveyed, however, the picture was slightly skewed towards larger organization sizes. Here too, while the largest proportion of organizations fell in the middle category (34.12% had between 100-500 technical and managerial employees), 22.35% had between 20 and 100 technical and managerial employees, and 23.53% had over 1000 technical and managerial employees.[75] Table 5.6 presents some of these key organizational structure variables.

[73] Bozeman et al. (1995) use the word "superlab" to denote the largest of the US National laboratories.

[74] About 45% had between 500-4999 employees, and 32.5% had between 5000-49,999 employees. The question regarding employment at the parent organization was worded to include all employees ("How many Employees does the organization have?" and asked the respondent to pick one of five non-linear categories, namely, <100, 100-499, 500-4999, 5000-49999, >50000. These were coded as 1 through 5 (in that order) in the database.

[75] The question regarding employment at the parent organization was worded to include only technical and managerial employees ("What is the best estimate of the # of technical and managerial employees at this

Our respondents (N=42), on average, spent 34% of their annual R&D budget in the central R&D labs, 49% in business unit labs, and 22% through external or contract research. They spent about 60% of their R&D budgets on basic and applied research (31% and 28% respectively), another 20% on new product development and process or product improvement (13% and 9% respectively) and about 17% on test and engineering and technical support combined. Nearly 40% of technical and managerial employees working in the responding organizations held a Ph.D., an MD, or an M.Phil. degree, about 21% held a masters degree, and 28% held a bachelors degree. Only 2.45% of the technical and managerial employees working for the responding organizations held a management degree (e.g. an MBA).

FIGURE 5.7: BREAKDOWN OF R&D EXPENDITURE & QUALIFICATION MIX (BY OWNERSHIP STRUCTURE)

Type of R&D Activity (Continuous%)	**Aggregate (N=79)**	**Private (N=27)**	**Public (N=23)**	**Academic (N=29)**
Basic Research	30.42	6.92	17.03	62.93
Applied Research	28.55	17.42	47.65	23.37
Product/Process Improvement	9.60	19.03	6.30	3.43
New Product Development	13.89	31.66	7.76	2.22
Test & Engineering	8.50	12.80	10.79	2.84
Technical Support	8.67	11.25	10.46	4.84
Educational Qualification of Technical and Managerial Employees (Continuous%)	**(N=77)**	**(N=27)**	**(N=23)**	**(N=27)**
Research Degrees (Ph.D./MD/M.Phil.)	40.36	25.69	33.34	61.02
Masters (MS/MSE etc.)	21.21	25.39	23.08	15.45
Bachelors (BS/BSE etc.)	28.03	38.58	33.79	12.58
Business (MBA)	2.5	3.46	1.26	2.59
Other Qualifications	6.91	7.79	8.51	4.67

The aggregate figures, however, must be taken with caution as they incorporate the biases introduced due to non-respondents. While some analyses may make sense with the aggregate data, others might require using one or more of the sub-categories to try to eliminate the effect of this bias. Looking at aggregate and sub-categorical statistics is useful, therefore, not only for identifying the direction and magnitude of the bias but also for the intrinsic motivation of wanting to learn

facility?" and asked the respondent to pick one of five non-linear categories, namely, <20, 20-100, 100-500, 500-1000, >1000. These were coded as 1 through 5 (in that order) in the database.

more about how organizations with different ownership structures may be different on various constructs of interest. We take up the latter issue in Section 5.2.

5.2.2—The Performance Profile of Responding R&D Organizations

The idea of organizational performance was central to our analysis, not only as an independent variable but also as a dependent variable[76]. Broadly speaking, there are two related, yet distinct notions of performance, namely, effectiveness and efficiency. Effectiveness relates to how well the R&D organization serves the mission of its clients. In the language of the scientific or technological community, that might translate into doing interesting work and being seen as a world leader by its peers etc. Efficiency, on the other hand, relates to the notion of productivity (or quality-adjusted productivity)—the number of patents granted or peer reviewed papers written per member of the technical staff or per unit of resource spent for the purpose. Effectiveness might be a predominantly qualitative or semi-quantitative measure of performance—for example, the perception of being useful to ones constituents or the scientific community at large— while efficiency might only be truly measured through quantitative means. Our survey tried to address both these notions of performance.

The first question dealt with a qualitative view of performance ("effectiveness"). The respondents were required to select one of four descriptive—yet distinct—categories that best described their "overall performance" with respect to their peers[77]. Due to the highly diversified nature of the respondents' R&D portfolios and the sectors in which they operated, it was deemed quite likely that multiple organizations may end up describing themselves as world leaders in their respect fields of endeavor. While there is a likelihood that self-described measures of performance, like the one we used, maybe biased, it may not necessarily be so. The results are presented in Table 5.8.

Of the 86 organizations that responded to this question, almost 28% described themselves as a "world leader" i.e. engaged in cutting edge R&D, another 17.5% describe themselves as being among the top-quartile or among the top-5 organizations in their class, almost a third (32.5%) describe themselves as above average, and another 22% describe described themselves as average (or below). While there is some hint of an upward drift here, the data is evenly distributed among high-performing (world leaders and top-quartile) and average or above-average performers.

---

[76] We were interested in performance of R&D organizations from both perspectives. For example, questions of the type: what are the characteristics of strategy-making, performance measurement, and incentive systems of high-performing vs. low-performing (or above-average vs. average) R&D organizations, use performance as an independent variable while those of the type: Does implementing a particular type of performance measurement system improve the performance of an organization use it as a dependent variable.

[77] The four categories were: a world leader (meaning engaged in cutting-edge R&D), in the top-quartile (i.e. among the top-5 facilities in its class, above average, and average (or other).

Whether this distribution represents some objective ranking of the responding organizations, we have no way to confirm without collecting additional data to corroborate these responses. The prima facie evidence, however, seems to suggest that the data does not suffer from the fallacy of the averages.

TABLE 5.8: KEY PERFORMANCE INDICATORS OF SURVEYED R&D FACILITIES

	Mean (Std) / (Range)
Qualitative	
Self-Rated Overall Performance *(1-4, N=86)*	2.51 (1.12) / (1-4)
Breakdown of Self-Rated Overall Performance:	
World Leader (engaged in cutting-edge R&D)	24 (27.9%)
Top-Quartile (~among the top-5 facilities in its class)	15 (17.44%)
Above Average	28 (32.5%)
Average (or Other)	19 (22.1%)
Quantitative:	
# of new product introductions in last 5 years (N=36)	78.13 (247) / (0-1125)
# of new projects initiated in last 5 years (N=49)	669.13 (1566.04) / (3-7000)
# of significant product improvements in last 5 years (N=32)	39.96 (77.98) / (0-300)
# of significant process improvements in last 5 years (N=27)	17.66 (N.A.) / (0-100)
# of new patents granted in last 5 years (N=65)	82.13 (146.77) / (0-1000)
# of peer reviewed papers published in last 5 years (N=50)	918 (1429) / (0-5110)
# of successful project transitions in last 5 years (N=22)	53.63 (108.58) / (0-459)

To get a handle on the constructs of productivity and efficiency, a comprehensive multi-item question required the respondents to provide quantitative and objective measures of their organization's performance. The question sought for the estimates of several indicators of R&D performance at the facility [in question] over the last 5-year period. We chose seven generic indicators of R&D performance that we thought were relevant to a wide range of R&D organizations and activities. These included: new product introductions, patents granted, peer reviewed papers published, product improvements, process improvements, and successful project transitions. Pelz et al. (1979) has found two of these indicators, namely, published papers and patents/applications, to be quite useful in identifying performance characteristics of individuals working in R&D organizations. The respondents were required to give an average value, aggregated over five years, of as many of these seven indicators as necessary or suggest one (or more) alternative performance indicators that were deemed critical at the R&D facility in question. We can use these measures of output in conjunction with measures of input (e.g. R&D expenditure etc.) to arrive at measures of efficiency and productivity.

In addition to the seven indicators provided to pick from, several of our respondents identified others as important indicators of their performance. These included: morale and satisfaction indices, research grants received, sales from new products, increases in stock price, staff retention rates, success in attracting federal funding, cooperative research and licensing activity, meeting schedule/cost milestones, and number of new ventures started etc. While the standard set of metrics we provided were aimed to establish a common baseline of performance among the respondents, the above measures reflect the diversity of their missions. These also illustrate the tendency among respondents to use intermediate measures of inventive activity (e.g. research grants obtained) as against final measures (e.g. papers published or patents filed as a result of those grants as a metric of performance.)

The data are clearly not very encouraging. The standard deviations around the means are huge—as much as twice the size (or more) of the mean itself. The means themselves are also seriously influenced by large outliers. Also, sample sizes are consistently smaller for each of the seven indicators, ranging from 65 for "number of new patents granted" to only 27 for "number of significant process improvements". In a later section of this chapter we will analyze these data for statistical correlation between performance and organizational structure variables of interest. For now, we would take this data as merely one of the several depictions of performance characteristics of our respondents.

5.2.3—Strategy-Focus and Strategic Themes Among R&D Organizations

We were interested in determining whether our respondents think about strategy as a source of "differentiation" in the marketplace of research and development activities or if there was homogeneity among the strategic postures of our respondents. The literature on traditional Balanced Scorecards suggests that organizations that do a good job of implementing a focused "differentiated" strategy tend to excel in one of the three strategic themes (i.e. technological leadership, customer intimacy, or operational excellence) while maintaining a minimum threshold of performance in the other two. While anecdotal accounts presented in Kaplan and Norton (2001) suggest this to be true for commercial entities and organizations, generally, we did not find any study that looks at this characteristic in R&D organizations, in particular.

In the first such investigation of its kind to our knowledge, we sought to determine how R&D organizations viewed strategy and whether or not they were similar to other commercial organizations in this respect. The respondents were given a choice of three possible strategic themes. They were also allowed to suggest an alternative theme, if deemed necessary. Respondents were required to pick as many strategic themes as they deemed appropriate—thus allowing us to assess

whether R&D organizations were trying to follow multiple strategic themes simultaneously. The results are presented in Table 5.9.

TABLE 5.9: TYPES OF DISTINCT TECHNOLOGY / R&D STRATEGIES

Technology / R&D Strategy (or Strategic Posture) (Prob, N=84)	**% of organizations**
Technical Excellence (Innovative Leadership)	738
Operational Excellence (On-time, On-Cost, On-Specs)	25
Customer Responsiveness (Internal and External)	571
Other	13
Number of Simultaneous R&D Strategies (N=82)	**# of organizations (%)**
# of Organizations having ONE R&D Strategy	40 (47.6)
# of Organizations having TWO Simultaneous R&D Strategies	25 (29.7)
# of Organizations having THREE Simultaneous R&D Strategies	13 (15.4)
# of Organizations having FOUR Simultaneous R&D Strategies	4 (4.7)
Average Number of Simultaneous R&D Strategies Employed	1.72

Of all respondents, 74% of the organizations identify "technical excellence" as the best descriptor of its strategic posture—a finding that is not surprising, given the population we are studying, although it might be biased slightly upwards for the simple reason that respondents in R&D labs might find it a little difficult to describe themselves as doing anything else but being technologically innovative. The proportion of organizations identifying operational excellence or customer satisfaction as their strategic posture were 25 and 57% respectively. Eleven organizations also identified a fourth strategic theme. These included strategic themes like: opportunities for students, quality and reliability, collaboration, workforce excellence, and inclusiveness etc.

A more interesting and useful finding pertains to the number of simultaneous R&D strategies (or strategic postures) adopted by the responding organizations. Our results indicate that R&D organizations are juggling multiple strategic themes simultaneously. Of all respondents, 47% described only one strategic theme as their primary focus, another 33% focused on two strategic themes simultaneously, and another 15% picked three themes simultaneously. On the whole, an average R&D organization in our sample implemented 1.72 simultaneous R&D strategies (or strategic themes.)

In order to assess whether organizations that maintain their strategic focus (a la Treacy and Wiersema, 1995) perform better than those that adopt multiple simultaneous strategic postures, we looked the strategic postures of sub-categories of organizations sorted on the basis of their self-described performance. The results are presented in Table 5.10.

TABLE 5.10: ARE HIGH PERFORMING R&D ORGANIZATIONS "STRATEGICALLY" MORE FOCUSED THAN THE REST?

Focus of Technology / R&D Strategy (%, N=84)	World Class	Top Quartile	Above Ave.	Average (Benchmark)
Proportion of Organizations Adopting a Strategic Theme	N=24	N=14	N=19	N=19
Technical Excellence (Innovative Leadership)	87* (2.67)	92* (2.67)	66 (.94)	52
Operational Excellence (On-time, On-Cost, On-Specs)	37 (1.58)	14 (.11)	25 (.80)	15
Customer Responsiveness (Internal and External)	66 (1.62)	57 (.83)	52 (1.13)	42
Other	17 (.08)	07 (.85)	11 (.60)	17
Number of Simultaneous R&D Strategies (N=82)	N=24	N=14	N=19	N=17
One Strategic Theme only	37	53	46	58
Two Simultaneous Strategic Themes	29	26	32	29
Three Simultaneous Strategic Themes	20	20	14	06
Four Simultaneous Strategic Themes	12	-	3	-
Average Number of Simultaneous Strategic Themes	2.08*(2.47)	1.71 (1.31)	1.62 (1.09)	1.35

Although the results are somewhat mixed, there are some significant findings from the above analysis. A significantly larger proportion of high performing R&D organizations (i.e. world class and top-quartile) describe technical excellence as their strategic posture as compared to average performers. The result is statistically significant at 5% level. An more important, and unexpected, finding is that, on the face value, the organizations in our sample seem to contradict Kaplan and Norton's hypothesis about strategic focus. High performing organizations, on average, pursue a higher number of distinct simultaneous strategic themes. For example, organizations that describe themselves as "world class", on average, pursue 2.08 simultaneous strategic themes as compared to 1.35 for those that describe themselves as "average" only. This difference is statistically significant at 5% significance level. While difference in the number of simultaneous strategic themes pursued by the top-quartile and above-average performers (1.71, and 1.62 respectively,) one the one hand, and average performers (1.35,) on the other, is not statistically significant, there is a clear trend in the direction of higher number of simultaneous strategies being related with a higher self-described performance rating. This is a very important finding—one that we will take up for discussion later.

5.2.4—Strategy and Performance Multi-Dimensionality Among R&D Organizations

In line with the Balanced Scorecard literature, we tried to assess whether R&D organizations took a multi-attribute view of performance and strategy. This also amounted to a test of the generic R&D

Balanced Scorecard presented in chapter-3[78]. We did this with the help of two questions (#16 and #17). The first question asked the respondent whether or not a particular dimension of performance was "considered critical" within his or her organization. The respondents were provided with the five performance dimensions (or perspectives) discussed earlier. The second question required the respondent to mark each of chosen dimensions against a five-point non-linear (descriptive) scale of successively increasingly managerial emphasis and control. The results are summarized in Table 5.11 and 5.12 below.

TABLE 5.11: PERFORMANCE DIMENSIONS DEEMED "CRITICAL" FOR R&D ORGANIZATIONS & DEGREE OF "BALANCE"

Performance Dimensions/Perspectives Deemed Critical (%, N=85)	**% of Organizations**
Employee Morale and Creativity (EMC)	71 (44)
Learning and Knowledge Management (LKM)	6 (49)
Innovation Management (IM)	57 (44)
Customer Satisfaction (CS)	64 (48)
Financial Performance and Control (FPC)	49 (50)
Other	25 (43)
	# of Dimensions
Average Number of Performance Dimensions Deemed Critical	3.06 (1.28)

Of the five pre-specified performance dimensions, employee and morale was the one considered critical by the largest proportion (71%) of responding organizations, followed by customer satisfaction (64%), learning and knowledge management (60%), innovation management (57%), and financial performance and control (49%). Several respondents also specified other performance dimensions as key to their overall performance. These included: safety and environment, student learning, new product introductions and management, technology disruptions, portfolio balance, external funding support, minimize defect rates, human subjects protection, consumer relevance, flexibility, stakeholders feedback, competitiveness vis-à-vis peer institutions, dealing with parent company, and job competence and performance.

On a broad level, although these data suggest that the underlying concepts behind each of these critical performance dimensions are not foreign to the respondents, the coverage is not as good as we would have ideally liked. Take the converse of these readings, for example. About 30% of R&D labs studied do not consider employee morale and creativity, a full 35% do not consider customer

[78] We would like to warn our readers, however, that a finding suggesting that R&D organizations strive to achieve balance between these five perspectives does not necessarily mean that they are implementing a Balanced Scorecard.

satisfaction, and about 40% do not consider innovation management as critical dimensions of performance. Additionally, the total number of performance dimensions deemed critical by a lab is about 3.06, on average, clearly indicating that R&D organizations, while taking a multi-dimensional view of performance, do not necessarily take a "balanced" view.

Another way to look at the same figures (not shown here) would be to look at the number and proportion of organizations in our sample who consider only one, or two, or three, or four, or all five performance dimensions as critical. About 10% of the organizations responding to our survey considered only one of the five performance dimensions as critical. Another 14% considered two of the five to be critical, 41% considered three of these to be critical, and 12.6% thought four of the five were critical. Only 18.4% of the organizations considered all five of the dimensions to be critical to their performance. This, we believe, is a solemn indictment of the "imbalanced" manner in which R&D organizations think about strategy and performance.

TABLE 5.12: HOW DILIGENTLY DO RDOS EMPHASIZE ON PERFORMANCE MULTI-DIMENSIONALITY IN THEIR STRATEGIC MANAGEMENT SYSTEMS?

Performance Dimensions/Perspectives (#)	Total # (std)	Max # (std)	Difference # (std) (Range)
Employee Morale and Creativity (EMC) N=61	2.85 (1.67)	3.59 (1.7)	.73 (.99) (0-4)
Learning & Knowledge Management (LKM) N=52	2.73 (1.69)	3.28 (1.7)	.55 (.84) (0-4)
Innovation Management (IM) N=52	3.62 (1.54)	4.15 (1.4)	.52 (.83) (0-3)
Customer Satisfaction (CS) N=56	2.85 (1.72)	3.33 (1.8)	.48 (.83) (0-4)
Financial Performance and Control (FPC) N=44	3.04 (1.39)	4.06 (1.4)	1.0 (.95) (0-3)
Other N=21	2.00 (1.89)	2.66 (2.3)	.66 (1.0) (0-4)

Considering a dimension as important or "critical" is one thing, and developing and implementing a strategic management system that executes on that piece of information and insight is another. In order to assess the "diligence" with which R&D organizations emphasized upon the critical performance dimensions, we developed a five-point non-linear scale to assess their commitment to integrating these dimensions in their strategic management systems (please see Figure 5.1 for the detailed item on the instrument).

The categories are devised in a manner that, for the strategic management system to be consistent, each higher category subsumes all lower categories. For example, unless an organization's management considers a particular performance dimension as important, it does not make sense to

emphasize it as a source of strategic advantage and unless performance in an area is explicitly measured and monitored, it makes little sense to try to tie rewards and incentives to performance in that area. If we find an organization that does the latter without doing the former, we have, in fact discovered an inconsistency in the strategic management system. The results are summarized in Table 5,12 (above).

17. Please mark which of the following statements are true for each of the five dimensions of performance identified in the above item (#16) <u>at this facility</u>.

DEGREE OF EMPHASIS PAID TO 5-KEY PERFORMANCE DIMENSIONS	EMPLOYEE (EMC)	LEARNING (LKM)	INNOVATION (IM)	CUSTOMER (CS)	FINANCIAL (FPC)
Considered important by management	☐	☐	☐	☐	☐
Management emphasizes it as a source of strategic advantage	☐	☐	☐	☐	☐
Explicit initiatives exist to develop a strategic edge in this area	☐	☐	☐	☐	☐
Performance in the area is regularly measured and monitored	☐	☐	☐	☐	☐
Rewards and incentives are tied to performance in the area	☐	☐	☐	☐	☐

Figure 5.1: Survey Item Assessing the Degree of Emphasis on Various Performance Dimensions

For employee morale and creativity dimension that was identified as critical to performance by 61 of 87 organizations, for example, the average degree of emphasis is 2.85 on a scale of 1-5. The average score on degree of emphasis variable for other performance dimensions range from the low of 2.73 and 2.85 in learning and knowledge management and customer satisfaction respectively to the high of 3.04 and 3.62 in financial performance and innovation management respectively. As expected, we found inconsistencies in the manner these performance dimensions were incorporated in the strategic management system. For almost all categories there was a 0.5 to 1.0 point difference between the total score and the maximum value of highest selection, suggesting loopholes in the strategic management systems. Many organizations claimed to do more (e.g. link rewards and incentives with performance) than they had the systems in place to for.

5.2.5—Performance Measurement Philosophy and Systems in R&D Organizations

One prevailing feeling within the R&D community is that R&D performance cannot be measured. One of the very well-respected expert-practitioner that we spoke to, prior to the development of the questionnaire, warned us that measurement of R&D performance was a "fundamentally wrong-

headed endeavor". We started our investigation into the performance measurement systems and practices by addressing this perception. We were interested in the perceptions of R&D leaders and managers about the potential uses of performance information within R&D organizations. We also sought to understand how performance-related information was actually used within these organizations. Comparing these two pieces of data would allow us assess the difference between perception and reality. The results are presented in Table 5.13.

TABLE 5.13: POTENTIAL & ACTUAL USES OF PERFORMANCE MEASUREMENT SYSTEMS/INFORMATION IN R&D ORGANIZATIONS

	# (Mean Value)
Perceived Potential Uses of Performance Measurement (N=81)	
Identify problems/deviations for action by mgmt.	3.91
Improve employee understanding of organizational systems	3.51
Align and communicate organizational objectives/goals	4.08
Provide input for performance-based rewards	3.76
Identify problems/deviations for action by employees	3.53
Justify existence, decisions, and performance	3.54
Motivate people by providing feedback	3.91
Actual Uses of Performance Measurement at the Facility (N=87)	**% of Organizations**
To provide a basis for incentives/recognition	65
To guide employee development policies	40
To provide feedback for employees/management	74
To aid in R&D strategy/policy-making	58
To communicate with corporate management	52

The results are quite instructive. There appears to be a general agreement on virtually all listed perceived (or potential) uses of performance measurement systems and information. With the exception of "align and communicate organizational objectives" that scores even higher, our respondents fall between agreeing and strongly agreeing on all other potential uses of performance information. These findings, however, only concern perceived or potential uses and not actual use of performance measurement systems within R&D organizations. The results on actual use are presented in the bottom panel of the Table 5.13. When comparing the figures on potential and actual use, we find a mixed picture. High support for strategic learning (more at the top-management level, and less so at the employee level) translates itself into a high degree of actual usage (among 74% of the organizations) for employee and management feedback. A mid-to-high support for usage to motivate, reward, recognize employees also translate itself into actual use in setting incentives/rewards /recognition systems among 65% of the respondents. The most significant

anomaly relates to the fact that a high support for usage in aligning and communicating organizational goals does not translate into actual usage to aid organizational strategy-making.

We also asked our respondents usage of specific performance measurement approaches. Among the responding organizations, the most commonly used performance measurement and management framework was Management-by-Objectives (MBO) that was used by 41% of the organizations. This was followed by Total Quality Management or TQM (used by 23% of the respondents), Balanced Scorecard (used by 20% of the organizations), and Six Sigma methodology (used by 17% of the organizations). There were some sectoral differences between public, private, and academic labs in the relative use of these methodologies.

TABLE 5.14: TYPES OF PERFORMANCE MEASUREMENT SYSTEMS OR FRAMEWORKS USED IN RDOS

Measurement Systems/Frameworks Used (%, N=85)	% of Organizations
Total Quality Mgmt.	23
Management by Objectives	41
Balanced Scorecard	20
Six-Sigma type Techniques	17
Technology-Value Pyramid	01
Other	17
	#
Total # of Frameworks Used	1.21
How Long in place (yrs.)	6.01
Perceived Impact	3.98

Several organizations used multiple performance measurement methodologies simultaneously. On average, an organization used 1.21 different performance measurement systems or framework at the same time and the average duration for the current performance measurement systems being place was about 6 years. On average, the organizations having a performance measurement system in place found them to have a "slightly positive" impact on performance within the organizations. The results are presented in Table 5.14.

We also asked our respondents to identify individual metrics or measurement approaches used to assess the performance of individuals as well as R&D facilities. The proportionsof organizations using a particular metric to measure individual or facility performance are presented in Table 5.15.

These metrics or methodological approaches are also ranked (1-5) in the order of the frequency of their usage.

TABLE 5.15: MOST USED (TOP-5) METRICS/MEASUREMENT APPROACHES FOR INDIVIDUALS AND OVERALL R&D FACILITIES

Metric or Measurement Approach (%, N=85)	For Individuals		For Facilities	
	%	Rank	%	Rank
Bibliometrics (e.g. # of patents, # of papers etc.)	.517	***2***	.4	***2***
Peer Assessment (e.g. external and internal reviews)	.588	***1***	.47	***1***
Professional Peer Recognition (e.g. awards and rankings)	.588	***1***		
Impact Metrics (e.g. # of paper/patent citations)	.329	***4***		
Measures of Teaching/Training Activities	.294	***5***	.23	
#/$ of Grants/Contracts obtained	.282		.32	***3***
Customer Satisfaction Measures	.270		.29	***4***
Improvements in Individual/Aggregate Capability	.270			
Technological Competence (e.g. # of PhDs, Post-Docs)	.235		.22	
New products/Services introduced in last x-years	.211		.28	***5***
# of Successful/unsuccessful Technology Transitions	.164		.24	
Quality of Strategic Goal Alignment and Attainment	.129		.21	
Self-described R&D Effectiveness Scores	.517	***3***		
Measures of Process Efficiency (e.g. success ratio of projects)	.588	***1***		
% Projects meeting Performance Specifications	.588	***1***		
Revenues/Cost Savings from New Products/Services	.329	***4***		
Employee Morale and Creativity Measures	.294	***5***		
ROI-type Measures (e.g. Yield, NPV)				
Time to Market (e.g. Development Cycle)				
Budgetary Growth Achieved			.25	

The results are self-explanatory. As expected, while peer assessment remains the most preferred measurement approach among both individuals and facilities, we find other differences between performance measures used to gauge individual performance vs. those used to measure organizational performance. Not only are the performance metrics or approaches different for individuals vs. facilities, the results are much more dispersed for facilities than they are for individuals. This highlights the very diverse spectrum of missions that R&D facilities are required to undertake. Not knowing the specific missions of the organizations we are trying to assess, we are not able to make a judgment on whether or not these are the right metrics to use—beyond making a passing judgment on whether or not using a particular metric, generally, makes sense for the subject in question.

5.2.6—Compensation, Incentives and Reward Systems in R&D Organizations

We were interested in the rules used for making compensation decisions within R&D organizations, especially, whether or not or to what extent individual or group performance seemed to matter in determining individuals' compensation. We also sought to find out whether these rules were applied for most employees or only a specific set of employees. Together these results would indicate the extent of a "pay for performance" culture within the organizations being studied. We also sought to determine the types of individual and lab-level incentives/benefits/rewards systems available to individuals at the R&D facilities. The results are summarized in Table 5.16 5.17, and 5.18.

TABLE 5.16 TYPES OF COMPENSATION "RULES" AVAILABLE AT THE FACILITY

Type of Compensation Scheme / "Rule" (%, N=85)	Types of Employees		
	Most Employees	Managerial Employees	Key Employees
Across the board % Cost of Living Increase	68	10	4
Years-in-service based Increase	28	3	4
Individual Performance-based Fixed Increase	28	9	11
Individual Performance-based Variable Incr.	54	27	31
Group Performance-based Fixed Increase	4	4	2
Group Performance-based Variable Increase	7	9	4
Total # of Simultaneous Compensation Options avail. (#, std)	1.95 (1.06)	0.67 (.88)	0.65 (.87)
Salary Ranking (self-described)	1.74	-	-

Table 5.16 describes the various compensation schemes or "rules" that are being used at the respondents' organizations for three categories of employees, namely, most employees, managerial employees only, and key-employees only. The results suggest that around two-thirds (68%) of R&D facilities adopt an across the board percentage cost-of-living increase-type compensation rule for most of their employees. This is closely followed by individual performance-based variable and fixed increases. About a third (around 28%) of the responding organizations also used a compensation system based on years-in-service for most of their employees—a system that has often been criticized in the literature as unfair and counterproductive to individuals' incentive to perform better.

The results also show a stark difference for managerial and key-employees category. For both managerial and key employees, the largest section of the responding organizations prefer individual performance based variable increases as a preferred rule for deciding compensation. None of the

other options had a significant prevalence. The results are in sharp contrast to an earlier finding in the literature that suggests that scientists and engineers prefer stability over flexibility and hence fixed over variable financial inducements. The respondents were allowed to pick as many compensation options under each category as were applicable to their organization and therefore it is likely that most organizations adopt multiple compensation rules to promote different kinds of performance behaviors from the same or different sets of employees.

TABLE 5.17: TYPES OF INDIVIDUAL-LEVEL INCENTIVES/REWARD OPTIONS AVAILABLE AT THE FACILITY (RESULTS FOR TOP-5 OPTIONS ARE HIGHLIGHTED)

Type of Incentive/Benefit/Reward Option (%, N=87)	**% of Organizations**
Time/money to attend professional events, conferences etc.	***78***
Share of royalty on patent/inventions	***58***
Cash rewards	***60***
Lab-wide recognition/token reward	***57***
Corporate-wide recognition/token reward	56
Fast track promotion within the technical track	51
Fixed cash reward per paper/patent etc.	31
Greater visibility among peers (invited to "important" meetings)	***60***
Access to discretionary funding (or time) to pursue interests	39
Creation of spin-offs or entrepreneurial sabbaticals	17
Ability to work on "prestigious" projects	21
Better independence/autonomy to pursue research agenda	28
Chances of being considered for research mgmt. position	21
Group-based fixed or variable performance rewards	21
Total # of Simultaneous Incentive Options Avail.	6.27 (3.09) (0-13)
Salary (self-described)	1.74

The data suggests that the number of simultaneous compensation options available within organizations is around 1.9. A quarter of the organizations (33%) have one compensation rule for all employees with another quarter (33%) having 2 compensation options, and about one-fifth (21%) having 3 compensation rules with the remaining 7% having 4 or 5 simultaneous compensation rules. The picture is different for managerial and key employees with much less diversity in compensation schemes. About half of the organizations (52%) of organizations do not have any additional compensation rules specific only to managerial or key employees over and above what is generally available for all employees.

Table 5.17 (above) presents the findings on individual-level incentives/benefits/reward options available at the responding R&D organizations. The results are self-explanatory. Responding labs use a mix of professional (or intrinsic) and financial (or extrinsic) rewards and incentives to incentivize performance by its individuals. The most used incentive (or reward) are professional (e.g. discretionary time and money to attend conferences, professional meetings, and skill-development activities) although financial rewards are also widely employed (i.e. in more than 50% of the responding organizations.) Those not found to be as widely applicable were group performance based awards (21%), entrepreneurial sabbaticals and possibility of creating spin-offs (17%), and improved chances of being considered for research management position (21%).

TABLE 5.18: TYPES OF LABORATORY-LEVEL INCENTIVES/BENEFITS/REWARDS OPTIONS AVAILABLE AT THE FACILITY (RESULTS TOP-5 OPTIONS ARE HIGHLIGHTED)

Type of Incentive/Benefit/Reward Option (%, N=87)	**% of Organizations**
Visibility of Lab and its employees	**55**
Ability to secure more funding in future	**51**
Corporate-wide recognition/token reward for the lab	**49**
Increased chances of promotion of lab's mgmt.	**31**
Improved ability to attract talent by the lab	**31**
Improved ability to negotiate better/interesting projects	22
Access by Lab's mgmt. to discretionary funding	**31**
Significant financial incentives for lab's management	10
Ability of Mgmt. to adopt attractive compensation schemes	05
Total # of Simultaneous Options avail.	2.88 (1.83) (0-7)
Salary (self-described)	1.74

Since the respondents were allowed the flexibility to choosing as many incentive options as were available at their organizations, majority of organizations chose several of them. On average, each organization had close to 6 incentives (or rewards) options available to its individual employees—a finding that suggested that organizations tended to develop a "menu of options" to suit the diverse needs and interests of its employees. This also supports a fairly robust finding in the literature that that organizations should develop flexibility in their incentive systems so that managers can reward every individual with the type of incentive chosen to best suit that individual's needs.

Table 5.18 (above) presents the findings on laboratory-level incentives/benefits/rewards options. More than any other option, (about 50% of) our respondents considered higher visibility of the lab and its employees (and its derivative effect on the employees' future employability and promotions) as the most common incentive available to the lab for performing better performance. Especially

striking, as compared to the corporate sector, was the absence of financial rewards for the lab's management as an incentive to improve the lab's performance. There might be several reasons for this finding, including but not limited to, concerns for equity as well as the understanding that an R&D manager has a very delicate role—not only as leader but also as a mentor to his colleagues and juniors—with little direct control over individuals' performance. Our findings appear to be consistent with the above point of view.

5.3—THE ANALYSIS OF SURVEY FINDINGS

While the above discussion has primarily focused on describing the various characteristics of R&D organizations in aggregate, we now turn to two key issues of concern to us. Are there significant differences between various features of strategy-formulation, performance measurement, and incentives systems of public, private, and academic R&D organizations? Also, are there any significant differences between above characteristics or better-performing vs. average R&D labs? Next, we address these issues.

5.3.1—Differences between Public, Private, and Academic R&D Organizations

The differences between public, private, and academic R&D organizations is an issue of central concern to this investigation. Do the ownership structures of an R&D lab make a difference in terms of how organizations conduct their strategy-formulation, performance measurement, and incentives design functions? Is there something germane to the way inventive activity is organized in each of these "sectors" that drives their performance effectiveness or efficiency? Table 5.19 presents a summary statistical snapshot of key variables of interest.

Before we discuss the findings, however, we must make a couple of clarifications. First, the public vs. private distinction has been an area of considerable debate in the performance literature for quite sometime now (e.g. Halachmi, 1996 and 1997, Crow and Bozeman, 1998, Bozeman, 1987.) Authors have contemplated the effect of a number of factors, namely, bureaucracy, multiplicity of objectives, non-definability of performance, concerns for accountability, lack of a clear bottom-line, and nature of missions etc. as reasons for differences in performance of public and private sector organizations (Osama et al., 2002).

Dr. Barry Bozeman of Georgia Institute of Technology, along with several of his colleagues, have defined "public-ness" as a characteristic of all organizations (Bozeman, 1987). According to this view being public is not a function of the legal ownership structure of an organization but a function of economic and political authority. While there is considerable merit in Bozeman et al.'s views on the subject we, for the purpose of this analysis, stick to the ownership/legal structure of organizations as a determinant of their public-ness. We are not claiming to do a causal analysis of the effect of legal structure of an organization on a set of dependent variables but rather an analysis

of how organizations under certain ownership/legal structures—driven, perhaps, by a third set of factors—may exhibit different behaviors and characteristics.

Second, our respondents clearly are not representative of the entire universe of R&D organizations. Neither was our sample frame a randomly selected one. Any findings, therefore, must be interpreted in view of the limitations that may result from the above. Much of the narrative below deals specifically with R&D organizations in our survey sample.

With that caveat in mind, the data does reveal several very interesting findings. We would start with the obvious and expected ones and then move on to some unexpected but very consequential results. First, private, public and academic labs differ in terms of the type of R&D work they perform with the substantial majority (~90%) of the work performed at academic labs being of research nature while the converse was true of private sector labs. Second, R&D organizations within private, public, and academic sector(s) differ in terms of their self-described performance ratings. The aggregate average self-described rating for all organizations was around 2.5 (falling midway between above-average and in the top- quartile range)—with the average for public sector labs being the highest (3.11) followed by private sector (2.39) and academic sector (2.03).

Several explanations may be advanced for the this finding, ranging from over-representation of high-performing public sector labs and under-performing private sector labs in our sample, to a possible "complacency" bias among public-sector lab managers. It is also possible that this may indeed represent the "true" assessment of competitiveness of the labs represented as well. Most of the public sector labs in our sample are one-of-their-kind defense, agency, and national laboratories with unique missions and they indeed are engaged in research and development at the cutting edge of their fields of endeavor hence qualifying them as among the top-quartile or world-class performers. With our data, though, we cannot isolate which of these explanations is valid.

The findings also suggest that public sector labs juggle a greater number of strategies simultaneously as compared to both overall and the private-sector labs. Not only do a greater proportion of public sector labs adopt each of the three strategic themes, they also adopt a greater number of strategic themes simultaneously (average of 2.24 as against 1.72 for the aggregate and 1.70 for private sector labs.,) These results are statistically significant at the 5% level. Academic labs, on the other hand, fall on the other side of the continuum. They are more likely to adopt technical excellence as a strategic theme and less likely to adopt customer responsiveness and operational excellence. On the whole, they have the least number of simultaneous strategic themes, on average—a result that is statistically significant at 10% level.

TABLE 5.19: DIFFERENCES BETWEEN UNIVERSITY, PUBLIC AND PRIVATE SECTOR R&D LABS ON KEY DIMENSIONS OF RELEVANCE (BENCHMARK: PRIVATE LAB)

Important Structural Variables	Aggregate (N=79)	Private (N=27)	Public (N=23)	Academic (N=29)
Self Described Performance Ranking	2.51	2.39	3.11*	2.03
Type of Lab (Research: Development)	59:41	21:79	61:39	90:10
R&D Strategies or Strategic Themes	%	%	%	%
Technical Excellence	73.8	64.2	76.9	80
Operational Excellence	25	28.5	38.4	10**
Customer Responsiveness	57.1	67.8	80.7	26.6*
Other R&D Strategy	13	7	20	13.3
Total # of Simultaneous Strategies	*1.72*	*1.70*	*2.24**	*1.3***
Degree of "Balance" in Performance Dimensions	%	%	%	%
Employee Morale & Creativity	71	57	84*	74
Learning and Knowledge Mgmt	60	35	57	83*
Innovation Mgmt.	57	75	46*	51**
Customer Satisfaction	64	71	92*	35*
Financial Performance/Control	49	5	53	45
Other Performance Dimension	25	42	19**	16*
Average # of Dimensions	3.29	3.32	3.53	3.06
Emphasis on Performance Dimensions	#	#	#	#
Employee Morale & Creativity	2.32	1.78	2.73**	2.5
Learning and Knowledge Mgmt	1.90	1.03	1.88**	2.78*
Innovation Mgmt.	2.64	3.39	2.34**	2.17*
Customer Satisfaction	2.07	2.14	3**	1.14*
Financial Performance/Control	1.79	2	1.84	1.53
Other	1.95	1.61	2	2.8
Performance Measurement Systems in Place	%	%	%	%
Total Quality Mgmt.	23	14	30	25
Management by Objectives	41	64	30*	29*
Balanced Scorecard	20	25	30	6*
Six-Sigma type Techniques	17	39	7*	6*
Technology-Value Pyramid	1	3	0	0
Other	17	21	19	12
Total # of Frameworks Used	1.21	1.67	1.19**	.80*
How Long in place (yrs.)	6.01	5.96	5.9	6.30
Perceived Impact	3.98	4	4.09	3.82
Compensation/Incentives Structures	#	#	#	#
Salary (self-described)	1.74	2.17	1.65*	1.39*
# of Simultaneous Compensation Options	1.90	1.85	2.26	1.64
# of Simultaneous Individual Incentive Options	6.27	6.07	6.57	6.19
# of Simultaneous Lab-level Incentive Options	2.85	2.64	3.15	2.80

* indicates a 5% significance level, ** indicates a 10% significance level, Private is the benchmark category

Public sector labs seemed to be much more "balanced"—as defined by simultaneous emphasis on several performance dimensions-- than either private sector or academic labs. While the results are not significant for the average number of dimensions considered "critical", several of the individual performance dimensions are statistically significant. For example, as compared to private sector labs, a statistically significant proportion of public sector labs consider employee morale and creativity, customer satisfaction, and innovation management as dimensions critical to their performance. Although academic labs seem to do worst in terms of the overall "balance", they too fare better than their private sector counterparts on several individual performance dimensions.

For the degree of emphasis on each of these performance dimensions—and the quality of their linkage to the strategic management system[79]—public sector labs seem to do better than private sector labs on all performance dimensions except innovation management (where they perform worse) and financial performance and control. Several of the results are statistically significant. While these results suggest a mixed picture, there is, nevertheless, a trend that one can distinguish in the data. Different sectors do better or worse on difference performance dimensions—and never better or worse overall. Public sector organizations seem to do well in employee and customer measures, private sector organizations in innovation management, and academic labs in learning and knowledge management. There is no statistically significant difference for financial performance and control. To a large extent, these results point at the historical and natural tendencies of organizations[80] to cater to varying organizational needs and missions. There is, however, a clear need for greater balance across multiple dimensions of performance among organizations belonging to all sectors.

Public sector R&D organizations also use, on average, lesser number of performance measurement frameworks than private sector organizations do. Average number of performance measurement frameworks in place per organization in our sample was 1.2 with highest for private sector (1.67) followed by public sector (1.19,) and academia (0.8.) The results are statistically significant for difference of means between public and private (at 10% level) and academic and private (at 5% level.) Among the types of frameworks used, private sector organizations are less likely to employ total quality management (result statistically not significant but suggestive) and more likely to employ six sigma type measures (result statistically significant at 5% level) than public or private labs. Balanced Scorecard is prevalent almost equally among public and private sector labs (30% and 25% respectively) although considerably less so among academic labs (6%). There was no statistically significant difference between the three sectors in terms of number of years since these

[79] We operationalize the quality of linkage between strategy, performance, and incentives systems, by measuring the degree to which performance multi-dimensionality thinking penetrates through the strategy-formulation, performance measurement, and incentives systems using a five-point descriptive scale (Q 17.)
[80] E.g. public sector's regard for multiple stakeholdership and equity, private sector's regard for efficiency, and academia's regard for innovation and learning

frameworks were in place, on average, or the perceived impact of these frameworks on organizational performance, although it seems to be that public sector labs seem to do slightly better than average and academic labs seem did slightly worse than average.
Finally, in terms of the compensation structure and incentive schemes, public sector organizations seem to lag behind their private sector counterparts. While private sector labs, on average, pay somewhere between 'equal or slightly above average' and 'consistently higher-than-average' salaries, public labs pay somewhere between 'below-average' and 'equal or slightly higher than average' salaries. Academic labs do even worse. These differences are statistically significant at the 5% level. Public sector labs, however, seem to compensate for the lower salaries through greater flexibility in compensation options and individual-level incentives options. Although the results are not statistically significant, they are, nevertheless, suggestive and consistent.

Public sector labs seem to do better than private sector labs on both counts, namely, compensation flexibility and incentive structure flexibility as measured by number of options available. Oddly though, public sector labs also seem to perceive stronger lab-level incentives for performance than academic labs or private sector labs. This is quite contrary to the general perception that bureaucracy and other factors force public sector lab managers (and their labs) to become complacent as they see little personal or lab-level benefit to them for working harder or performing better than before[81].

None of the other organizational features (e.g. large and small labs, basic-applied and development labs, and corporate and divisional labs returned practically significant differences on strategy, performance, and incentives dimensions (analysis not shown here). What does this imply for our analysis? While there are some statistically significant differences between large and small labs, research and development labs, corporate, divisional and overall company R&D, these are not as large and consistent as those between public, private and academic labs. Most of the statistically significant findings were either expected or can be explained without stretching the theory too much. A small sample size seriously hampers our ability to perform a multivariate analysis that controls for the observed differences. However, we interpret our results as an expression of the fact that, for our sample of R&D organizations, the public-private-academic distinction remains the most powerful of all distinctions that we have considered. Most certainly, there are several factors, perhaps unobservable in our data, that would account for differences among public-sector, private-sector or academic R&D organizations. To be fair, many of these differences emerge from underlying missions, requirements, beliefs, histories, and traditions of these organizations.

[81] For example, if one's research funding or budgetary allocation is to be determined through some kind of annual escalation formula rather than performance, one's incentives to perform may be seriously dampened.

This analysis is able to identify specific areas that need a more detailed consideration or ways in which organizational structures and thought patterns must evolve if we are to implement the Balanced Scorecard within each of these categories of organizations. Another view of the findings may be that the Balanced Scorecard needs to adapt to certain peculiarities of the various types of organizations to which it must be applied. If there is some virtue in the way R&D is organized within each of these sectors e.g. multi-stakeholdership within public sector and concern for efficiency within private sector—and, we believe, that to some extent there is—then care must be exercised in incorporating these special needs and virtues while implementing different elements of the Balanced Scorecard in R&D organizations.

5.3.3—Affect of Strategy-Performance-Incentives Systems on Organizational Performance

The second set of questions that are of interest to us relate to the effect of strategy-performance-incentives systems on organizational performance itself. The mail survey sought two types of performance information on the respondents. The first of these was the self-described performance rating that could be interpreted as a measure of technological ingenuity ("being world class or doing things at the cutting edge") and also as a proxy for effectiveness in achieving the mission of the sponsors. The second of these were a series of objective, although self-reported, performance metrics that the respondents chose to disclose as measures representing important determinants of their performance. These can be used to arrive at the notion of efficiency of R&D in these organizations.

While the results in quantitative measures of performance are inconclusive, primarily because of small sample sizes, we can glean some insights from the qualitative analysis. Table 5.20 presents one way of looking at the notion of performance. It looks at and compares various structural features (i.e. elements of strategy-performance-incentive systems) of organizations that described themselves as world-class and among the top-quartile (designated as high-performing,) on the one hand, and above-average or average (designated as low-performing,) on the other hand. The first thing we notice is the organizational breakdown of high-performing and low-performing organizations[82] in our sample, specifically, the predominance of academic or university labs in the former (50%) and public sector labs in the latter (51%) category. This raises the possibility that the differences between high-performing and low-performing R&D organizations thus obtained may be biased by a third variable, namely, their organizational structure rather than performance-rating itself.

[82] From here onwards, the terminology high- and low-performing merely refers to our own relative categorization of R&D organizations based on their self-described performance ratings. It is by no means a judgment on an individual lab relative to another or on several of them collectively.

TABLE 5.20: DIFFERENCES BETWEEN HIGH PERFORMING & ABOVE-AVERAGE/AVERAGE R&D LABS ON KEY DIMENSIONS OF RELEVANCE (DIFF. OF MEANS TEST)

Important Structural Variables	Aggregate (#, N=84)	High Performing (#, N=38)	Low-Performing (#, N=46)
Self Described Performance Ranking	2.51	3.63	1.58
Sector of Operation Private: Public: Academic	33:30.5:36.5	28:51:21	37:13:50
R&D Strategies or Strategic Themes	%	%	%
Technical Excellence	73.8	89*	60
Operational Excellence	25	28	21
Customer Responsiveness	57.1	63	52
Other R&D Strategy	13	13	13
Total # of Simultaneous Strategies	1.72	1.94*	1.52
Degree of "Balance" in Performance Dimensions	%	%	%
Employee Morale & Creativity	71	76	67
Learning and Knowledge Mgmt	60	53	65
Innovation Mgmt.	57	61	54
Customer Satisfaction	64	71	58
Financial Performance/Control	49	46	52
Other Performance Dimension	25	23	28
Average # of Dimensions	3.29	3.33	3.26
Emphasis on Performance Dimensions	#	#	#
Employee Morale & Creativity	2.32	2.37	2.28
Learning and Knowledge Mgmt	1.90	2.02	1.8
Innovation Mgmt.	2.64	3.13*	2.24
Customer Satisfaction	2.07	2.43	1.77
Financial Performance/Control	1.79	1.81	1.77
Other	1.95	2	1.9
Performance Measurement Systems in Place	%	%	%
Total Quality Mgmt.	23	33*	15
Management by Objectives	41	43	39
Balanced Scorecard	20	23	17
Six-Sigma type Techniques	17	17	17
Technology-Value Pyramid	1	0	2
Other	17	23	13
Total # of Frameworks Used	1.21	1.41	1.04
How Long in place (yrs.)	6.01	5.93	6.11
Perceived Impact	3.98	4.15**	3.81
Compensation and Incentives Structures	#	#	#
Salary (self-described)	1.74	1.89**	1.61
# of Simultaneous Compensation Options	1.90	2.05	1.78
# of Simultaneous Individual Incentive Options	6.27	6.58	6
# of Simultaneous Lab-level Incentive Options	2.85	2.64	3.04

* indicates a 5% significance level, ** indicates a 10% significance level, Private is the benchmark category

The more obvious and expected of the differences include the fact that a statistically significant number of high-performing (vs. low-performing) R&D organizations describe technical excellence as their preferred strategic theme. A greater proportion of high-performing organizations also describe operational excellence and customer satisfaction a preferred strategic theme—although the results are not statistically significant for either of the two. Also, high-performing R&D organizations pay higher salaries. The finding is statistically significant at 5% level. One unexpected and potentially consequential finding is that high-performing R&D organizations also had a higher number of simultaneous R&D strategies (or strategic themes). This result is statistically significant at 5% level.

Other findings of significance include the fact that high-performing R&D organizations do not statistically differ from low-performing ones in terms of performance multi-dimensionality on any of the five performance dimensions or in terms of balance across them. High-performing R&D organizations, however, paid greater emphasis on innovation management as compared to low-performing R&D organizations. The results were statistically significant at 5% level. Another significant finding relates to the use of performance measurement techniques and methodologies and their perceived impacts. A greater proportion of high-performing R&D organizations (vs. low-performing ones) used Total-Quality-Management (TQM) as a performance measurement methodology. None of the other techniques differed in terms of their prevalence among the high- and low-performing organizations. High-performing organizations also perceived the impact from the performance measurement approaches or techniques to be much greater than the low-performing ones. The results are statistically significant at 10% level. Also, high-performing organizations, on average, used a greater number of performance measurement methodologies or frameworks as compared to low-performing ones. This result is not statistically significant.

Finally, we did not find any statistically significant differences in incentives structures between high- and low-performing R&D organizations. Although high-performing R&D organizations seem to have a somewhat larger number of simultaneous compensation and individual-level incentive options and a somewhat small number of simultaneous lab-level incentive options available to them, neither of these results was found to be statistically significant.

5.4—Discussion of Key Findings

While the survey results indicate several interesting—even unexpected—findings, we will now highlight the ones of greater consequence to our central concern, namely, the appropriateness and application of the Balanced Scorecard to R&D organizations. We will discuss the broader implications of these findings—taken in conjunction with ones from the case studies and practitioner interviews—in some detail in chapter-8.

First, survey results clearly suggest that the use of the Balanced Scorecard as a performance measurement framework among R&D organizations may be more prevalent than often believed. Contrary to our expectations, we found the Balanced Scorecard (or some variation of it) as one of the top-three performance measurement philosophies employed by R&D organizations in our sample. It was used by 20% of our respondents, second only to Management-by-Objectives (41%) and Total-Quality-Management (23%), both of which pre-date the Balanced Scorecard movement by decades. The prevalence of the Balanced Scorecard was even higher if one considers only the private and public sector R&D organizations (25 and 30% respectively). This provides some evidence of the growing popularity of the Balanced Scorecard movement within public and private sector R&D organizations alike. We also found that while performance multi-dimensionality comes naturally to R&D organizations, they do not do a very good job at maintaining a balance across multiple dimensions of performance. Only 18% of the organizations consider all five of these dimensions to be critical to their performance, about 12% consider four out of five and another 40% consider only three of the five dimensions to be critical to their performance.

Second, another finding of significant consequence to this research is the manner in which R&D organizations think about strategy. The survey results suggest that R&D organizations tend to take a diffused rather than differentiated view of organizational strategy in that they simultaneously focus on multiple rather than a single strategic theme (e.g. technical excellence, customer responsiveness, or operational excellence.) We found that R&D organizations, on average, employ more than one (ave: 1.72) strategic theme (or focus) and organizations that describe themselves as world-class focus, on average, on 2.08 strategic themes. Kaplan and Norton (2001) would probably argue that these findings do not auger well for the R&D enterprise. Strategy is about making choices and having a game plan to differentiate yourself from your competition, asserts Michael Porter (Porter, 1996). Others might contend that R&D organizations are fundamentally different from commercial entities to whose needs the Balanced Scorecard is tailored to. Our findings may be interpreted as a prima facie evidence of the popular notion often expressed by NASA's expressed mission of being "Faster, Better, and Cheaper" at the same time. Recently, however, this mantra has increasingly come under public criticism (Oberg, 2000; Pencikowski, 2000.) This is one area that needs more detailed analysis.

Third, the survey results point out to significant discrepancies in the alignment between strategy-formulation, performance measurement, and incentives systems. We sought to address this by asking our respondents to identify the degree to which they emphasized upon each of their chosen "critical" dimensions of performance. This allowed us to ascertain the extent to which each performance dimension was taken into account while making organizational strategy, determining how its performance was measured, and how the employees were rewarded. We found a trend that

was uneven, at best. For all five dimensions of performance, we found discrepancies in the degree of emphasis throughout the strategy, performance, and incentives continuum. For example, while organizations claimed to reward individuals based on their performance in a certain dimension (e.g. innovation management,) they did not have a performance measurement system in place to support that activity. These discrepancies ranged from the highest for financial performance and control category to the lowest for customer satisfaction category.

Finally, the survey results point towards significant differences between public, private, and academic R&D organizations in key aspects of strategy-formulation, performance measurement, and incentives systems. We did not find any other organizational features such as size, type of R&D (e.g. research or development labs,) or organization of R&D activity (e.g. corporate or department labs,) to be a strong predictor of differences in strategy, performance, and incentives systems. To the extent that this is true, this might hint at a need for taking the cross-sectoral differences in the organization of R&D activity (e.g. mission requirements, resource constraints, ability to attract talent) into account in deploying the Balanced Scorecard in R&D organizations.

One major non-finding relates to the absence of statistically significant differences between high- and low-performing R&D organizations. We also did not find clear linkages between desirable characteristics of effective strategy-performance-incentives systems (e.g., use of multi-attribute frameworks, and quality of linkage between strategy, performance, and incentives systems) and improvements in organizational performance itself. We expect this to be an artifact of the data (i.e. lack of good measures of organizational performance) rather than the organizational reality itself. Small sample sizes and lack of important second-tier performance variables undermined our ability to explore possible reasons for these apparent anomalies. This lack of finding clearly calls for a more detailed examination.

On the whole, we are compelled to conclude that while we find some discrepancies surprises in the data, multi-attribute strategy and performance architectures, like the Balanced Scorecard, are not foreign to R&D organizations. This analysis represents an important first step toward our intended goal of developing a better sense of the universe as we make an informed judgment on the use the Balanced Scorecard in R&D settings. The obvious limitations of survey research coupled with problems of a low response rate do not allow us to draw detailed conclusions about how R&D organizations implement Balanced Scorecard and whether or not these make any impact on organizational performance itself. We will take up these issues in more detail in the case study segment of this research.

CHAPTER #6

STRATEGY-PERFORMANCE-INCENTIVES (SPI) SYSTEMS IN PRACTICE: LESSONS FROM INDIVIDUAL CASE STUDIES

Case studies were chosen to be one of the two methodological approaches—alongside a mail survey of R&D labs across the US—primarily because of their ability to illustrate the phenomena of interest, namely, strategy-performance architectures in R&D labs, with their complex and multi-dimensional contextual details within actual R&D settings. In this chapter, we would provide a brief snapshot of each of the six cases conducted for the purpose. For the more interested and diligent reader, we would recommend the more detailed case descriptions and analyses in Appendix A. In the chapter that follows, we would look at these cases from a methodological lens and do a cross case analysis to make generalizations across them.

Each of the case briefs follows a predictable pattern. We introduce the case study by highlighting its salient features with respect to strategy-making, performance measurement, and incentives systems. We then contrast the case with the traditional (Kaplan and Norton) and our own R&D Balanced Scorecard (developed earlier) and ask the question: How balanced are the strategy-performance-incentives systems of the organization in question? Does it implicitly or explicitly aspire for balance across performance dimensions or even follow the Balanced Scorecard? Does the strategy-performance architecture (or Balanced Scorecard) contain the necessary structural ingredients that deliver the sort of performance breakthroughs often attributed to the Balanced Scorecard? Finally, we conclude by briefly outlining the key lessons learnt from that particular case itself.

6.1—CASE STUDY #1:

A MULTI-PRONGED STRATEGY-PERFORMANCE-INCENTIVES SYSTEM AT A MEDICAL PRODUCTS CO.

CSO1 is a $50 mn. mid-sized spin-off of a multi-product international conglomerate specializing in medical, environmental, and consumer products segments. CSO1's is a traditional private sector R&D operation ($10 mn.) that is driven by the needs of the market. The 20 or so technical employees at CSO1 work in a fast-paced, results-oriented, project-driven environment that puts very high expectations of them with regards to the commercialization potential of their scientific and technological contributions. The nature of work, therefore, is focused on new product development (70%) and product and process improvement (25%) with basic research accounting for only 1% of

the R&D expenditure and applied research taking up the rest. The case illustrates several important aspects of strategy-performance-incentives (SPI) systems in private sector organizations relevant to our central thesis. For example:

- ***CSO1's strategy formulation system is fairly typical in its use of multiple constructs e.g. mission, vision, strategic objectives, values etc. in defining strategy*** that can, if they don't relate to each other, lead to confusion, dilution of focus, and "shallowing" of strategy among its employees**.**
- ***CSO1's primary performance measurement architecture—an MBO-based system—has its weaknesses and strengths that its leaders and managers must look into.*** While it does a fairly comprehensive job of cascading organizational strategy and objectives to the level of an individual employee, it can, at times, have limitations in terms of its ability to connect with the technical employees.
- ***CSO1's incentives structure emphasizes on giving its employees, including technical employees, a stake in the company's performance.*** While the incentive structures in place do not discriminate against technical employees, they also do not depend upon a clear and direct connection between their contribution or that of the R&D department to CSO1's financial performance.
- ***Although CSO1 does not implement a Balanced Scorecard, its strategy-performance-incentives systems aspires "balance", both explicitly and implicitly.*** The degree to which the balance is actually achieved varies considerably across performance dimension.

The defining feature of CSO1's strategy-performance-incentives (SPI) systems is the use of multiple managerial frameworks e.g. MBO, Six Sigma, 360-degrees, Standards of Leadership (SoL) model etc. alongside (but also on top of) each other in the strategy, performance, and incentive realms whose alignment presents a challenge to the organizational managers. The case raises critical questions about the marginal benefit/value derived from putting additional managerial frameworks in place (Please see p.6). Following is a brief overview of CSO1's strategy-performance-incentives (SPI) system that illustrates its key features and lessons learnt.

6.1.1—A plethora of sources contribute to a predominantly top-down strategy-making process— In a manner that is typical of small-to-mid sized private sector organizations, organizational strategy, priorities, and performance expectations are driven by the needs of the competitive marketplace. They are very precisely defined through consultations between the Board of Directors of the parent conglomerate and CSO1, the top management of CSO1, and the research management in a partly consultative and partly top-down manner. The individual scientists and engineers contribute to the process only indirectly—primarily as custodians of organizational scientific capability—but are responsible for on-the-ground implementation of R&D strategy. The real

challenge of the strategic planning process is to match organizational capability (the bottom-up stream) with external opportunity (the top-down stream). The annual strategic planning process begins with the identification of CSO1's profit and revenue targets. Organizational objectives are then identified in five key performance categories ("dimensions") of ***market leadership, growth, profitability, operational excellence, and people development*** designed to support the revenue/profit targets. This five-part categorization, however, differs from the performance perspectives of the Balanced Scorecard, not only in nomenclature and ordering but also in the lack of an explicit causal performance model behind the categories.

Several of these strategic objectives, spread across five performance categories, are non-financial (e.g. 14 out of 25 strategic objectives in a specific year were non-financial) and each relates to a specific deliverable for one of the several functional/departmental heads (i.e. Vice Presidents). These are termed OGSMs ("objectives-goals-standards-measures"). These highest-level OGSMs are then cascaded down the organization into OGSM2s (for directors) and OGSM3s (for managers and individuals). This process of cascading operationalizes the organization's overall objectives and strategy into pieces that are actionable by individuals (Figure 6-1). In principle, this system of cascaded OGSMs, by aligning the objectives of various sub-organizational units with that of CSO1, tightly couples the entire series of individual efforts into a coherent whole in a transparent and fair manner. In practice, however, we encountered some resentment by technical employees about the inability and inflexibility of this MBO-based system to account for and recognize their unique contributions to the organization.

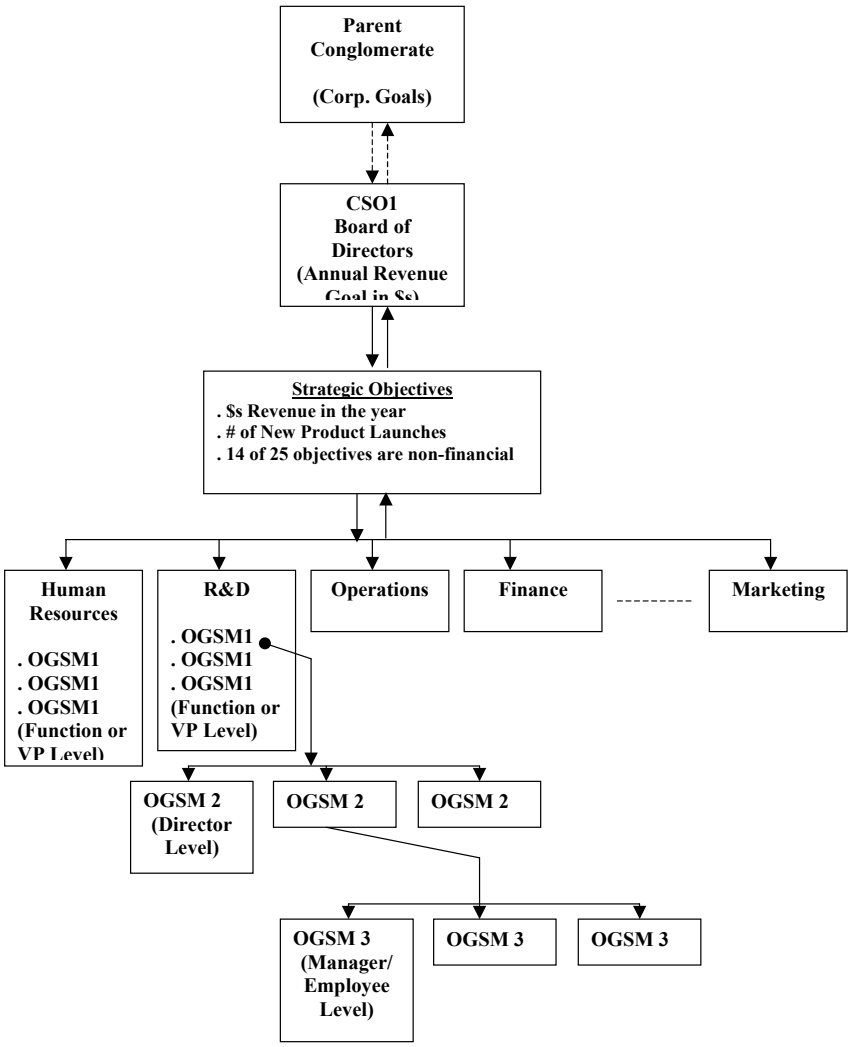


Figure 6-1: OGSMs— Hierarchical Strategic Planning at CSO1

Another possibility of dysfunction—of which we found some evidence—arises due to multiplicity of sources of guidance that impinge on the strategy formulation process. These include organization's mission, vision, values, and statement of objectives etc. Dysfunction may arise if these multiple sources of guidance do not follow through from (or relate to) one another or even if a case for it has not been made in an explicit and transparent manner. With the interrelationships between each of these sources (e.g. how do the constructs identified in the vision relate to those identified in the statement of objectives or vice versa) not clearly defined, the employees begin to consider them as separate and unrelated sources of guidance thus leading to organizational confusion and dilution of strategic focus.

6.1.2—A multi-pronged performance system measures strategy execution & organizational alignment—CSO1's performance measurement architecture presents a picture of an "over-managed" organization. The multi-pronged performance measurement system comprises several layers of measurement activity pieced together through at least three separate processes. The first (and main) framework in this series of measurement frameworks is the MBO-based OGSM system (discussed above). OGSMs for each level of the organization, starting from the departmental head all the way down to the level of an individual, are determined in a manner that identifies specific quantitative performance targets for each. These performance targets are then used to measure and monitor performance throughout the organization. A significant element of this process is the one-to-one translation of the five performance categories used for developing organizational strategy into ones used for measuring its performance as well.

The second of this series of performance frameworks—the SIPOC diagramming—has its roots in the Six Sigma methodology. Through this framework, CSO1 maps over a 100 important processes across varied organizational domains, identifies performance benchmarks, and determines data requirements for measuring the performance of each of these processes. SIPOC diagramming is carried out every 2 years or so and the alignment between diagrammed and actual processes is checked every year through an organization-wide Process Excellence Audit. The performance measures and targets identified for each of the processes whose SIPOC diagram is created are also back-linked to organizational objectives.

The third of this series of frameworks is an individual performance appraisal system administered through the HR function within CSO1. This system, in principal, integrates well with the OGSM-based performance architecture by directly adopting the individual level OGSM3s as measures of individual performance but also adding upon them. The system uses a Standards of Leadership (SOL) model of assessing individuals' development needs whose connection and contribution to

organizational strategy is, at best, only speculated. In addition to the above, CSO1 also plans to implement a 360-degree feedback system to further strengthen its performance architecture.

Both SIPOC and individual performance appraisal (OGSM) systems seem to have some, albeit non-quantifiable, impact on performance itself. The mechanism for such an impact, however, was haphazard rather than systematic. For example, the individual performance appraisal process—through its quarterly employee-supervisor meetings—forced people to work towards their goals. The usefulness of these meetings, however, largely varied from individual to individual. Some managers-employees paid more emphasis to making these a useful way of stocktaking and performance feedback while others treated them as merely an exercise in paper-pushing. SIPOC diagramming and the Process Audit, one the other hand, helped culminate into a check of organizational alignment and correction of deficiencies in organizational processes with significant opportunities for performance enhancement. This process, however, happened only once every two-years with the consequence that it did not provide for a sustained focus on organizational strategy and performance.

6.1.3—Incentives structure emphasizes making employees stakeholders in organization's success—CSO1 provides a fairly high-powered incentive system and competitive environment for its employees—the focus primarily being to provide the technical employees with a stake in the company's success. Both financial and professional rewards exist for technical employees. The key vehicle for professional rewards/recognition is the dual career ladder that rewards individuals who excel within the technical domain. Titles such as "Distinguished Member of Technical Staff" and "Chief Engineer or Scientist" are fairly common in the industry and are established as standards for recognizing performance of technical professionals.

On the financial side, there exist performance-based salary increases, an annual bonus plan for all employees (including technical employees), and a stock options plan for key executive-level employees. Annual raises are awarded from an annual merit pool and differential raises based on individuals' performance are common. Bonuses are structured in a manner to incorporate individual, departmental, and company-wide performance. Generally, individuals, based on their performance, can get anything from 8% to 40% (on average) of their basic salary as an annual bonus—although this figure may exceed well above the 40% mark for key managerial employees. Stock options are available to key employees as a means of executive compensation that, based on meeting and exceeding of performance goals, might go anything up to 100% of base salary or more.

According to the prevailing corporate wisdom, financial incentives are important for incentivizing desired performance by the employees, including the technical employees. The use of OGSMs to assess an individual's performance tightly couples individual motivations to organizational goals

and provides them with an incentive to outperform these goals. CSO1's financial incentive structures are designed to provide its employees with a stake in the company's success. The merit pool, the bonus plan, and the stock options plan are designed with that objective in mind. CSO1, however, must strive to ensure that rewards and recognition of the technical staff is as closely aligned to their own performance as possible and watch for the situation where it becomes too dependent on the financial success of the company over which they have little direct control.

6.1.4—Though not a Balanced Scorecard, CSO1's strategy and performance systems aspire for "balance" across performance dimensions—Although CSO1 does not implement a Balanced Scorecard (a la Kaplan and Norton, 1996, 2001 etc.), it does strive to achieve balance across multiple dimensions of performance in its strategy and performance measurement systems. Figure 6-2 presents a picture of CSO1's strategy-performance multi-dimensionality laid over the performance perspectives of the traditional Balanced Scorecard (in dotted gray). Despite significant differences in the specifics, the two frameworks are, in general, quite similar.

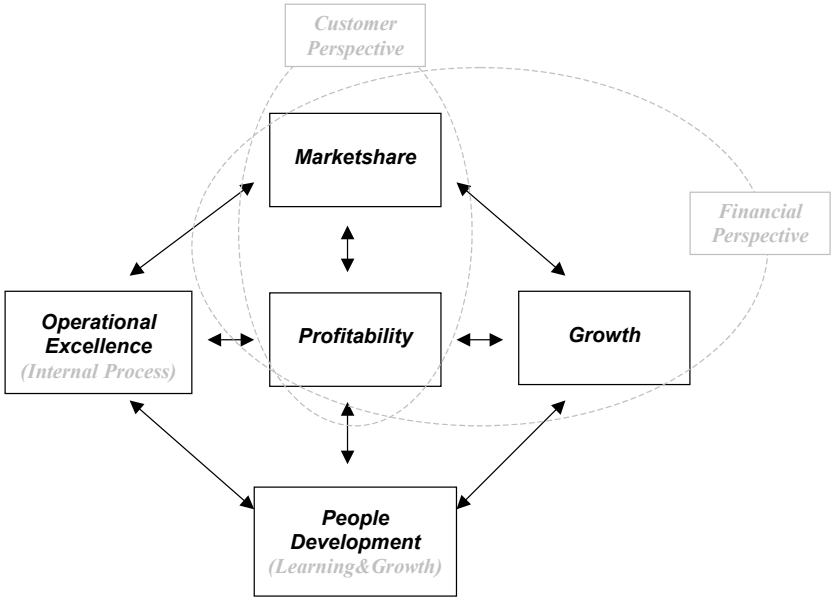


Figure 6-2: Balance in Strategy-Performance Systems at CSO1

CSO1's strategy-performance-incentives system also indicate several of the structural features that one finds in the Balanced Scorecard and similar systems, including, performance multi-dimensionality, cascading and operationalization of strategy, its use for setting rewards and incentives, and its ability to integrate with the broader strategic management framework etc. On several other counts CSO1's strategy-performance-incentives system appears to fall short of the standard practice among Balanced Scorecard adoptees, namely, the explicit use of cause-and-effect logic, the participatory nature of strategy-making process, and the use of actual scorecards and strategy-maps etc.

Figure 6--3 presents another view of CSO1's strategy-performance-incentives systems. Here we compare the explicit (assessed through a simplified strategy question: "Is this performance dimension deemed critical") and implicit (assessed through a detailed strategy questionnaire)

degree of balance in CSO1's SPI systems using our own R&D Balanced Scorecard as a benchmark. The results indicate that although CSO1 does not implement a Balanced Scorecard, it may explicitly or implicitly aspire for balance across multiple dimensions of performance. For example, even though CSO1's leaders and managers did not believe they emphasized customer satisfaction to be a critical performance dimension, a detailed strategy questionnaire indicates that they did do most of the things one is likely to find in an organization that does so.

6.1.5—Do Multiple Management Frameworks Constitute a Coherent Strategic Management System—Does the "MBO-based" system of management at work at CSO1 provide a comprehensive and integrated strategic management framework that sends an unambiguous signal to the employees? In principle, it can, although in practice, it probably falls short of the ideal. The use of multiple management frameworks at CSO1 presents clear evidence of the fact. We, and several of our interviewees, struggled with the notions of "marginal" benefit CSO1 was getting out of these successive layers of management tools and frameworks. Do CSO1's leaders gather a different set of metrics for their OGSMs and a separate set for SIPOC diagrams? It is likely that already over-burdened and skeptical employees see these as apparently disjoint measurement frameworks that do not necessarily link well with each other. The competing pressure of having to pay attention to each of these may also lead to measurement "overload", confusion and lack of focus among CSO1's employees.

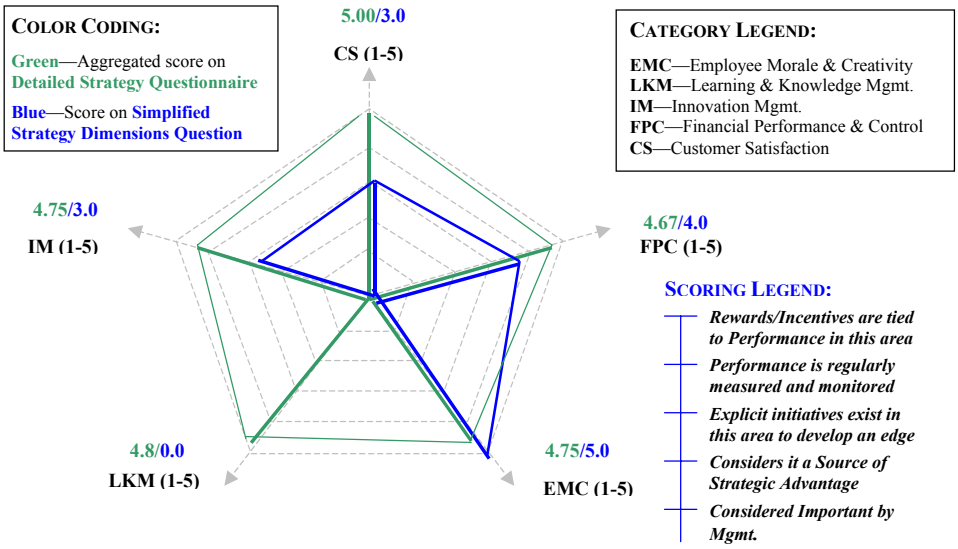


Figure 6-3: Strategy-Performance-Incentives Alignment at CSO1

Also unclear is the additional marginal benefit of having these multiple performance measurement frameworks in place or, conversely, if they represent an unwarranted "measurement burden" or drain on CSO1's resources. An official of CSO1, for example, failed to come up with a clear case for marginal benefits from multiple techniques, defending the use of SIPOC diagramming as a requirement set by the parent that may or may not be generating additional value at CSO1. We found some evidence of "measurement burden" as well. For example, much of the SIPOC

diagramming and data-collection activity happens not as an on-going exercise to ensure organizational alignment but as a one-off thrust right before the Process Excellence Audit.

6.2—CASE STUDY #2:

PARTICIPATIVE STRATEGY-MAKING AND PERFORMANCE DASHBOARDS IN ACADEMIA

CSO2 is mid-sized state-funded land grant institution of higher education with an annual R&D expenditure of $120 million. The university hires around a 1000 employees in technical and managerial/administrative cadre and can be subdivided organizationally into 7 constituent colleges comprising around 45 academic departments. In terms of performance, it compares rather unfavorably on several measures of research input, process, and outputs among a list of 12 peer institutions. The case illustrates several important aspects of strategy-performance-incentives (SPI) systems in academia that are relevant to our central thesis. For example:

- ***CSO2's compact planning process represents the cutting edge in participative strategic planning within academia and research settings.*** Though hindered, initially, due to non-uniform understanding (and hence implementation) of the underlying concepts, compact planning has the potential to involve and empower researchers and align sub-organizational objectives with that of the larger organization.
- ***CSO2 uses dual performance measurement frameworks, namely, traditional peer-review/tenure system and performance dashboards, in line with the needs of its academic and administrative functions respectively.*** The challenge obviously for CSO2's management is to ensure that this dual system does not put competing demands on individuals and sub-organizational units as well as the need to bring about a gradual convergence in the dual architecture.
- ***CSO2's incentives structure is fairly traditional for bureaucratic and equity reasons*** and the administrative and research leadership emphasizes intrinsic motivation and better upfront selection of (researchers) employees as keys to individual performance.
- ***CSO2's performance dashboards resemble the Balanced Scorecard in many ways but they also differ in several significant respects.*** While there is an explicit desire to achieve balance, the degree to which it is actually achieved varies considerably across various sub-organizational entities.

The distinguishing feature of CSO2's strategy-performance-incentives (SPI) systems, however, remains its use of cutting edge strategy-making and performance measurement methodologies, namely, compact planning and performance dashboards, in academia—a practice in which contemporary academic research institutions lag far behind their industrial counterparts. This case illustrates some of the problems associated with implementing such "in-vogue" management

frameworks in traditional academic environment and hints at the possibilities of performance enhancement that might arise out of these.

Following is a brief overview of CSO2's strategy-performance-incentives (SPI) system that illustrates its key features and lessons learnt.

6.2.1—Compact planning represents the cutting edge in participative strategy-making in academia—The Compact Planning process, although quite new and innovative, is not entirely new to academia. It has been inspired by its success at several well-known institutions of higher learning, namely, North Carolina State and Texas A&M University, among others. At CSO2, the concept of compact planning formalizes a series of iterative and participative top-down and bottom-up consultations that go into the annual strategic planning process. The idea, because of its emphasis on a promise or a contract (a "compact") between two parties, in this case academics/researchers and the administration of CSO2, is especially suited to the academic settings because of their tradition of academic freedom, flexibility, and independence.

In theory, a compact at CSO2 is a *formal agreement of a pre-specified format, written jointly by the contracting parties, that delineates directions and actions, respective responsibilities, investments, outcomes, and mutual performance expectations in the context of unit and university's long-range goals.* The compact planning process, if correctly undertaken, provides a means for involvement of reluctant academic and research staff into the strategy-making process, encourages-forces them to think through the organizational priorities and strategy and their own contributions to it, and come up with strategies (along with budgets, performance targets, milestones etc.) to bridge the gaps. Starting from the top and repeated throughout the organizational hierarchy, compact planning provides a means to align sub-organizational priorities with organizational goals and helps identify opportunities to rationalize investments and performance enhancement activities.

In practice, however, implementation details might vary as do the perceived or realized benefits. At CSO2, for example, compact planning has been mandated throughout the organizational hierarchy, including administrative and academic units/colleges. While the administrative units find it quite easy and "natural" to implement a management framework, the academic units have found it challenging to do so. The most important of these challenges include uneven understanding of the basic concept among academic/research staff who are required to develop the compact plans and skepticism about its value and/or the seriousness of CSO2's management. The first of these has led to faulty implementation while the second has led to lack of involvement in the process itself. Another major challenge has been to restructure or integrate the more traditional ("ivory tower") departmental peer review into a more commonplace compact planning exercise.

While the acceptance among the academic units, as against the administrative units, has been slow to begin with, even among those that have taken it seriously, the quality of implementation and the perception of benefits (and hence objectives of doing it) have varied considerably. One of the college deans we talked to could clearly relate instances where some of the department heads were quick to get the "knack" of the process and used it intelligently to get favorable budgetary allocations while others struggled to relate to it and found themselves at the short end of the stick. The "grumbling in the ranks", however, is slowly giving way to seriousness as the skeptical realize the benefits they stand to gain if they "play along" intelligently. The key factor here in building support for the process has been its connection to the annual budgeting process that unequivocally communicates the degree of management commitment to the entire organization.

6.2.2—Dual performance measurement frameworks put competing demands on CSO2's employees—The performance measurement architecture at CSO2 represents a mix of the new and the more traditional. CSO2's top-level leadership emphasizes the use of performance dashboards to display and communicate performance information to external, and more recently internal, stakeholders. The performance dashboards at CSO2 resemble the Balanced Scorecards in terms of their visual ("look and feel") and categorical features (i.e. performance dimensions). The use of these dashboards as scorecards of CSO2's performance for external audiences, however, contrasts sharply with that of the Balanced Scorecard that were first used for internal audiences only. To that effect, performance dashboards at CSO2 have largely been used as a "ceremonial" public relations rather than a strategy-mapping tool. While performance dashboards often follow organizational strategy, this correspondence is far from being perfect. There is, however, an effort underway to closely link performance dashboards with the compact planning process at CSO2 so as to fully ground it in organizational strategy.

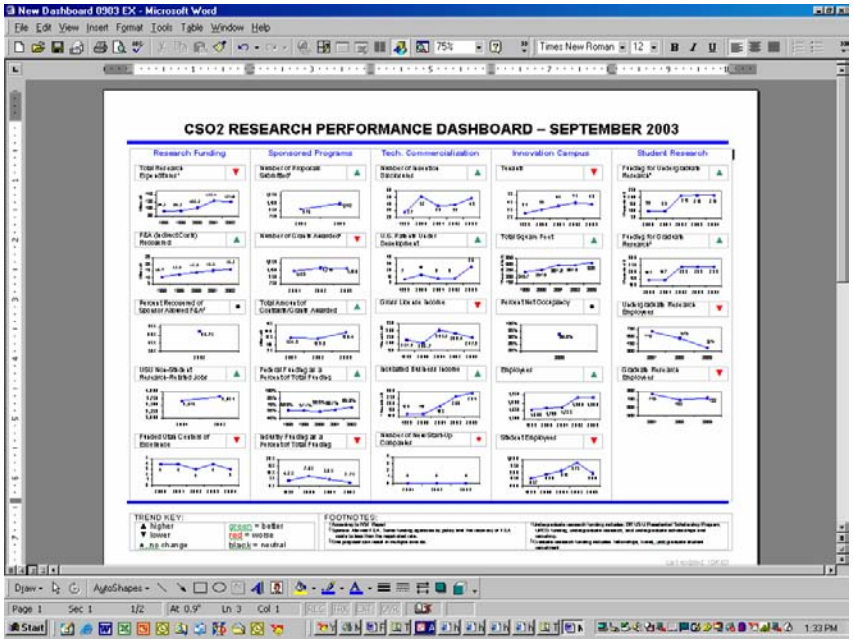


Figure 6-4: Performance Dashboards at CSO2

The adoption of performance dashboards has thus far been rather sporadic and primarily limited to the administrative units of CSO2. In the academic departments, on the other hand, the more traditional peer review/tenure system still remains the prevalent mechanism of performance

assessment. This process takes place in a collegial manner with each individual academic/researcher meeting with his or her departmental superiors and/or members of the tenure committee appointed by the Dean to discuss performance expectations as well as actual performance. These expectations may take the form of quantitative (e.g. score above a threshold on student reviews or number of published papers every year) or qualitative (e.g. contribution to intellectual environment of the institution) performance targets. The departmental peer-review/tenure system has been a fairly well-established practice in academia and is known for its strong influence on individuals' performance.

The two above-mentioned systems—found prevailing in different sub-organizational units—can, in principle, continue to function seamlessly if they emphasize the same performance attributes or even if they stress upon different aspects of performance. In practice, however, CSO2 possibly runs into a situation when the two conflict with each other, putting competing demands on individuals they seek to influence. For example, technology transfer--an organizational priority, as identified by the goals of the university--appears on one of the performance measurement framework (i.e. the dashboards) but not the other (i.e. departmental peer-review). Many researchers are compelled by the contradictory pressures to do technology transfer to contribute to their departments' performance dashboards, on the one hand, and to focus their energies on teaching and publication to satisfy the peer-review/tenure process, on the other. CSO2's leadership must work towards removing this apparent discrepancy or incorporate both within the same framework with explicit weights assigned to each so as to send an unambiguous message through the performance measurement system.

6.2.3—Intrinsic incentives and better upfront selection are perceived as key to performance—The rewards and incentives system at CSO2 is quite traditional, to say the least, mainly depending upon intrinsic factors (e.g. pride in one's work, flexibility and freedom etc.) to motivate performance. The incentive environment on the administrative side is low-powered with most of the officers/employees working on fixed salaries. This, however, does not rule out some strong reputational factors (e.g. chances of landing a better job at a different institution or promotion within the same institution) that might influence individuals' performance. Increasingly, the administrative and support services of the university are being converted into business-like operations with leadership's desire to unambiguously take the university in the direction of a stronger coupling between pay and performance. Several factors, namely, tradition, bureaucracy, and concerns of equity currently hinder the progress on this end.

On the academic side as well, CSO2 faces a constrained incentives environment. The strongest incentives are hypothesized to be intrinsic and professional i.e. not financial. A constrained resource situation and bureaucratic rules leave CSO2's leadership with little leeway in using financial rewards to incentivize the performance of teaching and research staff. Promotion to higher levels

(i.e. the tenure process) and greater prestige in the professional community that comes with it is by far the strongest motivator of performance. The prevailing wisdom seems to be that financial rewards, over and above the "carrot" of the tenure system, can do little to force people to work smarter or become better researchers. If given additional financial resources to improve performance, the institution would probably spend it on paying higher salaries to attract better researchers in the first place rather than trying to "force" getting additional output out of the existing ones. While true to an extent, we do also find some emphasis on and effect of professional-cum-financial rewards (e.g. startup funding, travel and conference support, recognition, re-assignment of duties etc.) on performance.

6.2.4—Performance dashboards incorporate several features of the Balanced Scorecard, but misses others—CSO2's performance dashboards resemble the Balanced Scorecard (a la Kaplan and Norton, 1996, 2001 etc.) in certain aspects of their structure, if not the intent. It identifies several dimensions of performance in its strategy and performance measurement systems. Figure 6-5 presents a picture of multi-dimensionality in CSO2's performance dashboards laid over the performance perspectives of the traditional Balanced Scorecard (in dotted gray). Despite significant differences in the specifics, the two frameworks are quite similar, generally. One important element that is missing in CSO2's performance dashboard, however, is the internal process dimension.

Figure 6-VI presents another view of CSO2's strategy-performance-incentives systems. Here we compare the explicit (assessed through a simplified strategy question: "Is this performance dimension deemed critical") and implicit (assessed through a detailed strategy questionnaire) degree of balance in CSO2's SPI systems using our own R&D Balanced Scorecard as a benchmark. We see a picture of an organization that is fairly well-balanced across performance dimensions although the degree and depth of alignment between strategy-performance-incentives systems falls short of the ideal.

From a structural standpoint, apart from balance between various performance dimensions, several other features stand out in CSO2's strategy-performance-incentives systems. These include, among others, the participative nature of the compact planning process,

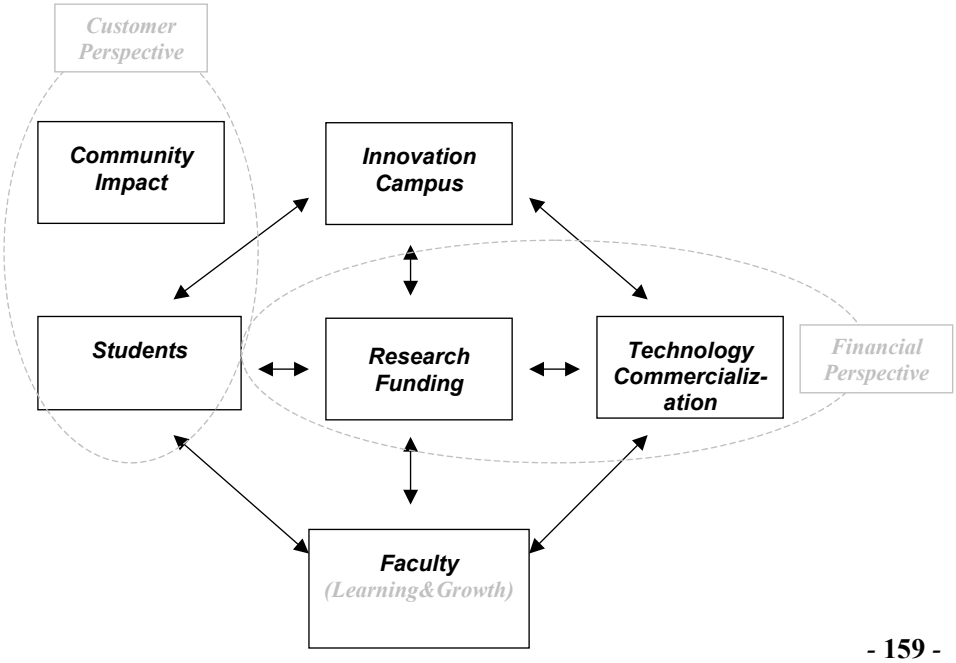


Figure 6-5: Balance in Strategy-Performance Systems at CSO2

the cascading of compact planning and performance dashboards throughout the organizational hierarchy, their ability to integrate (or integrate with) other aspects of the overall strategic management framework (e.g. budgetary process, individual performance appraisal system etc.), and the visual and communicational aspects of the scorecards etc.

Where CSO2's SPI systems lack are the absence of an explicit cause-and-effect logic (thus depriving it of the ability of strategic learning), the absence of a critical performance dimension, namely, internal process on the scorecard, and the lack of use as an internal strategy-making tool. Another serious deficiency in the CSO2's SPI systems is a less-than-desired connection between strategy (compact planning) and performance measurement (dashboards). These and other elements must be adequately addressed by CSO2's leaders and managers to fully exploit the potential of these innovative management frameworks.

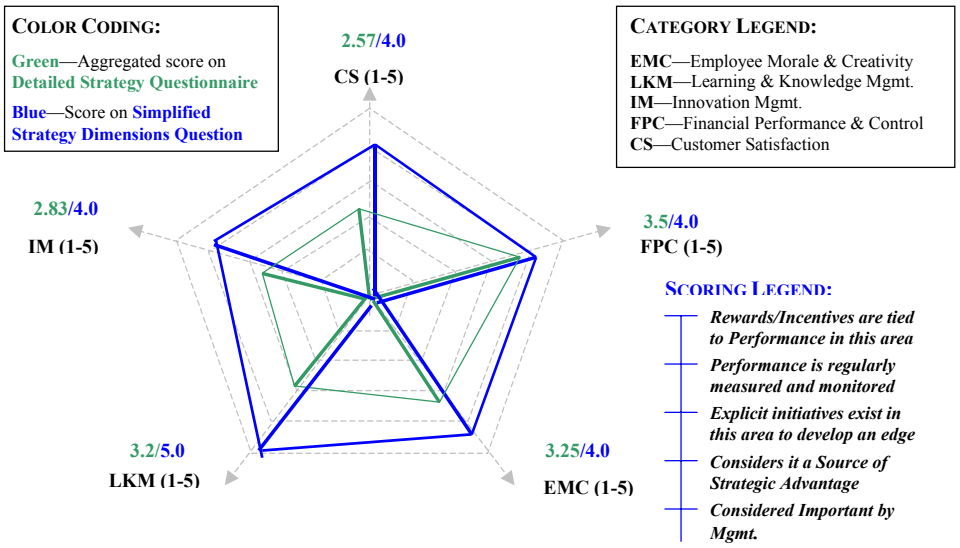


Figure 6-6: Strategy-Performance-Incentives Alignment at CSO2

6.2.5—Innovative strategy and measurement frameworks, if implemented intelligently, maybe well-received in academic/research settings—While CSO2's is a clear example of the efficacy of innovative managerial frameworks in more traditional academic research settings, it also illustrates several issues that might arise during such exercises. Compact planning process stands out as an example of participative strategy-making aimed at getting skeptical researchers and academics involved in the exercise. Top management support--CSO2's President's unequivocal support and involvement in the process—was critical to transforming skeptical bye-standers into eager participants. It also helped to raise the stakes of adopting or not adopting the process. CSO2 did so by linking its budgetary allocation process with sub-organizational compact plans.

Another important ingredient of a successful implementation, especially within the research settings, is a phased approach to implementation. At CSO2, the implementation of both compact planning and performance dashboards started with administrative units. The more skeptical and "conservative" research units only adopted these methodologies once a clear value proposition was

established for them. Persistence came out to be another important factor that first convinced the skeptics to "play along" and ultimately become enthusiasts. Persistence, however, is a double-edged sword that must be used with great caution. At least a couple of important obstacles also stand out:

- Firstly, the case established the importance of ensuring that everybody not only understands the tool/methodology being implemented but also its purpose within the organization in as clear terms as possible. At CSO2, we observed differences in enthusiasm for and success of the compact planning exercise, across various organizational sub-units. The organizational leaders, however, worked patiently with those units that did not do well the first time to ensure that they understood the process well and used it intelligently the next time around.

- Secondly, the case establishes the need for phasing in the newly adopted management technique in a manner that is non-threatening to the existing strategy-measurement architecture. This can be done by entirely replacing the old architecture with the new one—a difficult task by all counts, or by making the new complimentary to the old—especially as long as significant stakes remain are attached to the older way of doing things. In the context of CSO2, we can see this tension at work. Compact planning is gradually replacing the departmental reviews (a desirable end-state) while the performance dashboards are still complimentary and playing second fiddle to departmental peer-reviews in important ways (an intermediate state, at best).

Despite the above obstacles, however, CSO2 stands out as a proof of concept of several of the hypothesized advantages of Balanced Scorecard-type SPI systems in academic settings.

6.3—CASE STUDY #3:

A SIMPLE STRATEGY-PERFORMANCE SYSTEM AT A LARGE CORPORATE R&D ORGANIZATION

Case study organization # 3 (hereby, CSO3) is the North American arm of the central research and development organization of a large multi-company international conglomerate specializing in electrical and electronics, automation, transportation, and communication systems. As one of the five laboratories in the overall central global corporate R&D organization of the parent conglomerate—CSO3 conducts application-motivated basic research and advanced development in computer and communications technology and related areas. From an organizational standpoint, CSO3 is subdivided into two sub-organizational units, namely, the research and the technology labs. The research lab has an annual budget of $8 million that is divided between basic and applied research in the ratio 3:1. CSO3 employs a staff of over 50 professionals. The case illustrates several important aspects of strategy-performance-incentives (SPI) systems at a fast-dying breed of Bell-Labs-style R&D organizations that have to balance the dual missions of

curiosity-driven basic research and application-specific applied research within the same organizational entity. The contradictory requirements of both these realms of activity, namely, doing unfettered research at the cutting-edge while also justifying relevance to the corporate parent, creates interesting problems of strategy, performance, and incentives at CSO3 that are relevant to our central thesis. For example:

- ***CSO3's strategic planning process represents a series of one-to-one interactions between the research director and autonomous principal investigators/project leaders*** and is typical of this class of R&D organizations. The role of the research director, in this case, is central to the effectiveness of the organizational arrangement. He or she is a dual-purpose gate-keeper, balancing and mediating between the knowledge of research capability on the one hand, and corporate strategy on the other.
- ***CSO3 uses a traditional industrial peer-review type performance measurement system whose beauty (and effectiveness) lies in its simplicity.*** The challenge, however, is whether or not and how well does that simple performance measurement framework capture the complex reality of the organization and its relationship with the corporate parent/sponsor(s).
- ***CSO3's incentives structure is fairly traditional—depending upon intrinsic factors—although it does recognize extra-ordinary effort.*** The flexibility to hire and fire along with strong performance expectations and a tight and transparent feedback loop provides a potent (dis)incentive to perform
- ***CSO3's strategy-performance architecture although far from being a Balanced Scorecard, does exhibit elements of balance across performance dimensions.*** CSO3 also raises important questions about the effectiveness and/or marginal utility of using a Balanced Scorecard in small organizations with simple and uni-dimensional missions.

The distinguishing feature of CSO3's strategy-performance-incentives (SPI) system, however, is the simplicity with which it defines performance expectations, measures and rewards performance, and ties all of these together into a seamless whole. A key lesson to be drawn here relates to the need for communicating performance expectations as clearly as possible—even before an individual is hired into an organization—and then ensuring that all organizational systems send a coherent message to the employees by fully supporting the achievement of these expectations. From the standpoint of R&D leaders and managers alike, "hiring right", connecting with employees, and providing them with individualized feedback and incentives is critical to performance. While the specifics of the case might only be relevant to smaller R&D operations in less bureaucratic environments, the general points about performance expectations and feedback, nonetheless, apply to all R&D settings irrespective of size or type of work performed

Following is a brief overview of CSO3's strategy-performance-incentives (SPI) system that illustrates its key features and lessons learnt.

6.3.1—A highly individualized strategy-making process balances individual autonomy/freedom with corporate interests—CSO3 is organized as a relatively flat organization with the real authority and resources lying with individual scientists and engineers who work independently and autonomously as principal investigators or as members of small project teams. The strategy-making process at CSO3 attempts to, one the hand, mediate between its dual objectives of *"generating highly significant intellectual contributions"* (the scientific mandate) and *"significantly impacting parent's business [through transfer of technology]"* (the organizational mandate), on the other hand, the personal interests and agendas of a group of high-caliber scientists and engineers whom it hires and gives a relatively free reign over resources and intellectual agendas.

This process starts even before a formal strategy session is held. Specifically, performance expectations are clearly defined at the time of hiring of each and every individual within the organization. This is done through formal and informal discussions that explicitly touch upon what the organization stands for, what it expects from each of its employees, in what terms each will be measured, and what are the consequences of not performing up to those expectations. New hires are strongly encouraged to opt into the organization if their own personal intellectual agendas match that of the organization and if they are personally comfortable with having to take on the dual duties of innovating at the cutting edge as well as transferring technology back to the parent company. "Hiring right" in the first place, according to CSO3's managers, saves a lot of trouble later on when people could other find themselves out of sync with the culture and objectives of the organization.

The organization-wide strategy formulation process takes the form of a series of one-to-one consultations between research directors and principal investigators (for CSO3 Research Lab) or project leaders (for CSO3 Technology Lab). The research directors play a pivotal role in this exercise, trying to mediate between and align the intellectual agendas of the individual researchers, on the one hand, and long and short-term interests of the corporate parent, on the other. This interaction may take several forms and iterations. Lab directors, having listened to principal investigators' research priorities and plans for the coming year, may suggest minor modifications to make them more aligned to corporate goals, enhance chances of a successful technology transfer, advise on alternate approaches to achieving the stated objectives, or bring other related fruitful opportunities to their attention. Whatever the nature of these conversations, the decisions are made through mutual consent and understanding of the two parties.

The research director's unique position as a participant in the strategy-making echelons at multiple levels, namely, at CSO3 itself, at other labs in the parent's global R&D Network, and at the corporate parent, enables him or her to facilitate this process. CSO3 also appoints a dedicated executive—a

Vice President—as a liaison between CSO3 and the corporate parent. The above process plays out, with some variations, across both the Research and the Technology Labs at CSO3.

6.3.2—A mix of an industrial peer-review and MBO-based performance measurement system works at CSO3:

Performance measurement at CSO3 is relatively simple. Its beauty lies in the ability of CSO3's management to consolidate the entire set of performance expectations from its technical employees into a couple of performance measures. These are: quality adjusted publication activity (representing technical excellence) and major or minor technology transfer actions (representing relevance to the corporate parent). The publication and technology transfer activity of each individual is integrated over several years to avoid year-to-year fluctuations. Performance reviews are held every 4-5 years to assess each employee's technical merit and contribution to the corporate parent—at which time promotion decisions are generally made. The process is akin to a tenure process at a typical university, albeit modified to the industrial/corporate setting.

FIGURE 7: STANDARDS OF PERFORMANCE AT CSO3

Job Title	Years of Service at CSO3 / post Ph.D.	Scientific Achievement	Tech.-Transfer Record	Leadership
Research Associate	*0/N.A.*	*B.S., M.S. or equivalent experience*	*N.A.*	*N.A.*
Research Scientist	*0/0*	*Ph.D. or equivalent research experience*	*N.A.*	*N.A.*
Senior Research Scientist	*2+/6+*	*15+ peer reviewed papers of all kinds, 3+ major papers; some external recognition OR equivalent product/ business development experience*	*3+ minor tech. transfer activities e.g. h/w, s/w insertion in CSO3 product; procurement of funding; or licensing of Intellectual property*	*Key role(s) in 1+ group projects*
Distinguished Research Scientist	*6+/12+*	*30+ peer reviewed papers of all kinds, 6+ major papers; significant external recognition OR equivalent product / business development experience*	*1+ major tech. transfer activity e.g. initiation of new or significantly improved product; influencing strategic direction of R&D; significant licensing of IP.*	*1+ group projects initiated and managed*
Research Fellow	*8+/15+*	*35+ peer reviewed papers of all kinds,*	*One major tech. transfer success*	*2+ group projects initiated and*

Figure 6-7: Performance Expectations at CSO3

The effectiveness of CSO3's performance assessment system hinges on two factors. First, standards of performance for successive levels of seniority within the organization are well-defined and communicated. Individuals at different seniority levels are assessed and rewarded (or promoted) against the standards. Figure 6-7 presents this information in a tabular format. Second, the manner in which CSO3 applies this relatively simple performance measurement approach consistently throughout the organization has as much to do with its success as the simplicity of the performance framework itself (we will revisit this issue in section

6.3.5). The simplicity and consistency of the framework creates considerable pressure on each and every employee to perform. The intense pressure to perform as per a clearly defined set of expectations has, several times in the past, forced non-performing employees to voluntarily resign even before he or she has become due for promotion or re-evaluation.

To be fair, however, there are some criticisms of the system. The most common of these being that the approach emphasizes publications and technology transfer in a manner that is disproportionate to the amount of effort required for each. Technology transfer requires factors far beyond the technical expertise of a scientist or engineer to succeed and may put the outcomes of his or her efforts beyond his or her own control. This is valid criticism that needs some attention.

6.3.3—CSO3's traditional incentives structure depend on intrinsic and professional factors for motivation—The incentives system at CSO3—like other organizations of its type—is fairly traditional. The ability to decide one's own research agenda and work on projects that fire one's intellectual appetite is perhaps the most powerful incentive CSO3 provides to its employees. While salary raises and a bonus pool do seem to reward performance, the incentives structure generally depends upon professional rewards and recognition. One significant piece of financial reward—one that has never been awarded in CSO3's relatively short history—is designed for individuals engaging in extraordinarily heroic efforts that show the promise of considerable benefit to the company. These "extraordinary" awards are to be made for technology transfer efforts that significantly distract an individual from his or her research for a prolonged period of time, e.g. a year—thus emphasizing CSO3's commitment to technology transfer activity.

On the professional side too, there are several ways to reward performance, some intrinsic, others partly extrinsic—the most important, perhaps, being the opportunity and resources to do challenging work. According to company documents, the most promising projects get the most intern slots, consultants, and equipment. The support for a research project also depends on the (promise of) performance of the key employees on that project. In addition to research support and the greater flexibility in choosing one's research agenda, titles, promotions, and public recognition of one's achievement also serve as professional incentives.

In addition to the positive incentives, CSO3 has an aggressive negative incentive regime that strongly discourages below par performance. If a person fails to show the ability to meet the performance criteria of his or her present position—either due to lack of competence, or aptitude or interest towards a particular aspect of work, e.g. creating value for the corporation—he or she is expected to either fill-up the performance deficit or leave. This mix of positive and negative incentive climate makes CSO3's incentive structure a credible and potent influencer of individuals' performance.

6.3.4—Strategy-performance architecture, though simplistic, exhibits several attributes of the Balanced Scorecard—CSO3's strategy and performance architecture emphasizes two dimensions of performance, namely, technical excellence and technology transfer efforts. One may possibly argue that these dimensions of performance emphasize output at the expense of process and input measures and are hence are not "balanced" on the input-process-output-outcome continuum. That criticism would be overly simplistic. While CSO3's SPI architecture does not explicitly emphasize balance between performance categories, it does do so within a particular category. For example, it makes a careful use of process and output measures within the technology-transfer domain. In the initial stages of a scientist's career, technology transfer performance is measured by involvement in technology transfer activity (e.g. # of projects initiated) which is replaced later on by the output measures (e.g. creating a profit making product/service for the company) as one progresses higher into the organization and becomes more responsible and resourceful.

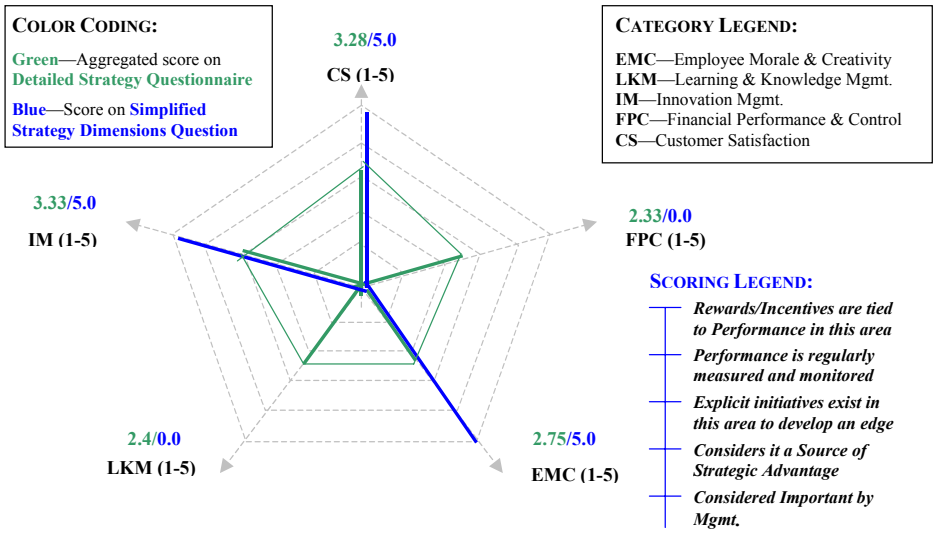


Figure 6-8: Balance in Strategy-Performance-Incentives at CSO3

Would having separate categories for internal process and learning and knowledge management make any difference for CSO3? It might, in principle, though the small size of CSO3's operation makes a tempting case in favor of a more informal approach to performance measurement. Figure 6-8 presents a snapshot of the degree of balance achieved by CSO3 within its informal environment. One can clearly see that CSO3's leaders take a "balanced" view of the organization at least implicitly, if not explicitly. The relatively low scores of detailed strategy questions (in green) probably refer to the lack of formal performance measurement in several sub-dimensions of performance—a fact that small organizations can probably get away with much more easily than their larger counterparts.

CSO3's strategy-performance-incentives architecture also lacks several of the structural features of the Balanced Scorecard (e.g. explicit use of certain performance dimensions, the cause-and-effect

model etc.), does not address several others because it may not require them (e.g. cascading), and addresses several others in a manner that differs from a traditional Balanced Scorecard (e.g. it operationalizes the strategy through the one-to-one interactions between technical staff and research directors). In this respect, barring the few weaknesses (discussed above), CSO3's strategy-performance architecture does seem to provide the kind of benefits one would expect from a Balanced Scorecard. The small size of CSO3's operations enables it to gain from several of the benefits without the requisite structural paraphernalia of the Balanced Scorecard.

6.5.5—The key to the effectiveness of CSO3's SPI systems lie in its simplicity and communication—CSO3's strategy, performance and incentives systems, barring a few issues identified above, seem to be in considerable harmony with each other. Employees participate in a one-to-one strategy-making process that is designed to bring individual goals closer to organizational objectives. Performance expectations and standards are clearly and simply defined in terms of two competing as well as complementary attributes, namely, technical excellence and technology transfer. Everything that an individual does must contribute to one of these two expectations. Everything else is either secondary (to the extent that it only contributes partially to the two primary objectives) or not important at all. This helps CSO3's management and employees in focusing their energies towards a small number of critical activities. Their performance is clearly measured according to these pre-set standards. Individuals are later rewarded for meeting or exceeding these standards.

Several factors contribute to the simplicity and elegance of this arrangement. The small size and simple structure (e.g. flat organization) of CSO3 makes it easy to communicate and manage performance expectations. Another simplifying factor is the ability of the lab's leadership to negotiate its mission/performance expectations with the sponsor. Equally important is to reduce is to reduce these expectations into a small number of performance metrics that can be easily measured and monitored. CSO3 exhibits a customized, individualistic style of research management. Research managers treat the work of each and every technical employee as a unique individual and works with him or her to help them deliver on their potential. The research director would, at times, even learn about his employees' work expertise to be able to provide the kind of mentoring needed to help him or her perform and regularly attend conferences and talk to experts or his/her colleagues to gauge his/her performance. As the organizational size increases, it would become more and more difficult to implement this system of individualized management and mentoring.

Similarly, it is not impossible to visualize a situation where CSO3's simple bi-dimensional performance measurement system would not suffice and a more complex set of performance measures would be needed to fully capture the performance expectations (for example, at a larger, more complex, R&D organization with several constituencies e.g. various funding agencies,

legislatures, etc.). In fact, it is perhaps quite natural for CSO3 to "grow out" of its simple yet effective SPI system as it evolves into a larger more complex organization. For example, CSO3 has so far resisted the temptation to compete for Federal R&D contracts even if they would help support its mission towards its corporate parent. It is logical to expect that allowing for that possibility would require additional measures to safeguard against the possibility of diverting from its mission thus complicating its simple SPI architecture.

Most importantly, however, CSO3 illustrates the value of effectively and unambiguously communicating organizational strategy and performance expectations to all employees. This case illustrates the powerful idea that a simple framework, even though it may not capture all the dimensions of performance, if communicated well, may be more effective in influencing performance than a more elaborate one that captures all but falls short in terms of employees understanding of what is required of them and how their performance is measured.

6.4—CASE STUDY #4:

IMPLEMENTING A BALANCED SCORECARD IN AN R&D FACILITY IN A PUBLIC SECTOR AGENCY

Case study organization # 4 (hereby, CSO4) is an R&D site of a national public sector agency charged with scientific research, technology development, and testing in aviation safety, quiet aircraft technology, small aircraft transportation, and aerospace vehicles system technology. With over 3800 employees—two-thirds of whom are civil servants—and $700 million in R&D expenditure, CSO4 is a premier R&D facility of its type in the US. The case illustrates several important aspects of strategy-performance-incentives (SPI) systems at public sector R&D organizations that are of importance to our central thesis, specifically, the use of a Balanced Scorecard at such a facility. For example:

- ***CSO4's strategic planning process is highly organized and well-documented with multiple sources of guidance available to make strategy.*** The resultant strategies, however, are fairly stable over time, partly representing long R&D cycles and partly the bureaucratic nature of the process itself. CSO4's challenge, therefore, is to put together a strategic planning process that is cognizant of its slow and bureaucratic techno-political environment but also nimble enough to affect real and meaningful change in its strategy when one is required.
- ***CSO4 implements a Balanced Scorecard that directly follows from a key source of guidance in the strategic planning framework.*** CSO4's Balanced Scorecard—designed as a Stakeholders' Scorecard—emphasizes the satisfaction of its customers (internal and external), funders (e.g. congress), and employees as three critical dimensions of its success.
- ***CSO4 has a fairly elaborate rewards and incentives system in place that focuses, primarily though not exclusively, on professional rewards and recognition to motivate individual***

performance. Constrained by civil-service-regulations, CSO4's leadership compensates for its lack of incentive intensity by providing a "menu of incentives". CSO4's leaders, however, describe intrinsic factors ("a higher calling") as the key motivators.

- ***CSO4's Scorecard, although balanced along a key dimension, namely, the interests of organization's various stakeholders, nonetheless, falls short of balance along other important attributes*** (e.g. input, process, output, outcome measures or present vs. future performance etc.). CSO4's implementation of the Balance Scorecard also misses several critical structural features of a Balanced Scorecard implementation thus limiting the benefits that could potentially accrue from it.

Despite its shortcomings, the distinguishing feature of CSO4's strategy-performance-incentives (SPI) systems, is that it is perhaps one of the most serious first attempts at modifying and implementing a Balanced Scorecard at a major public sector R&D organization and represents a proof-of-concept of the same. This case illustrates some of the problems and challenges associated with such an undertaking—most important of which is the challenge of integrating such a performance measurement framework in the generally bureaucratic and top-down strategic planning system so as to enable its true potential for empowering the employees.

Following is a brief overview of CSO4's strategy-performance-incentives (SPI) system that illustrates its key features and lessons learnt.

6.4.1—A highly organized strategic planning system draws upon multiple sources of guidance— The corporate vision statement of CSO4's parent agency emphasizes the unique role it has been charged with by the US Congress. The strategic planning process at CSO4 is highly organized, to the extent of being bureaucratic. Several sources of strategic guidance impinge on the process itself. CSO4 identifies three "critical success factors" (CSFs) that are instrumental in helping it achieve its objectives. These include, product users (i.e. customers), funders (i.e. sponsors), and CSO4's organization (i.e. employees and systems). A

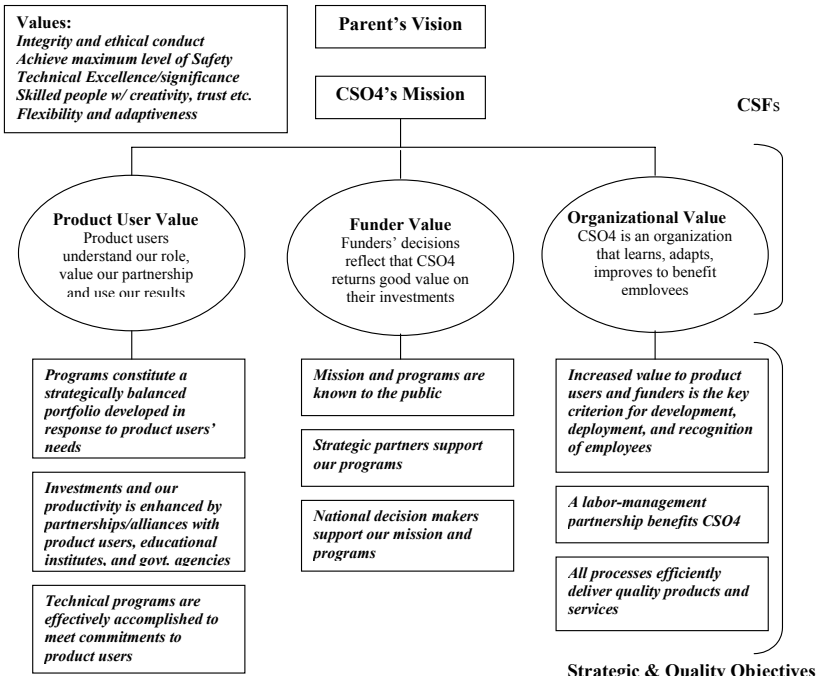


Figure 6-9: CSFs & SQF—Strategic Planning Architecture at CSO4

strategic and quality framework (SQF) that identifies specific objectives within the three-tiered management framework (described above) drives strategic decision-making at CSO4.

The strategic planning process at COS4 comprises well-documented series of steps aimed at aligning CSO4's activities with those prescribed in its mission, primary processes, critical success factors, and strategic and quality objectives. Objectives for each organizational unit (OU)—a standard sub-organizational entity at CSO4—are documented in an Organizational Unit Plan (OUP) reviewed and approved by the Director of CSO4. These OUPs also get guidance from CSO4's Mission Statement, a Thrusts Plan, and an Implementation Plan. Together, These four documents form an elaborate strategic management system with the OUPs functioning as the official mechanism for the definition and approval of ground-level organizational changes within CSO4's OUs.

The process begins, around January of each year, with an off-site retreat of the Center's top-leadership and unit managers, to discuss center performance, future directions, and actions and steps needed to implement these. This process has traditionally been quite inflexible in terms of accommodating drastic changes in organizational strategy from year to year—a fact that can partly be attributed to the bureaucratic nature of the process and partly to the long R&D cycle times in the aerospace industry. Consequently, it is viewed skeptically by the scientific community as an activity that takes time and energy and delivers little value in return. That, however, is beginning to change as more scientists and engineers are getting involved in this process.

6.4.2—CSO4's Balanced Scorecard ties performance measurement with strategy through the SQF— CSO4's performance scorecard has evolved from the traditional Balanced Scorecard 6-8 years ago. It uses the same set of performance categories deemed critical in the strategy formulation process and thus directly follows from the latter. The individual metrics on the scorecard are decided through a process of trial-and-

CSO4 Center Metrics

Product User Value

1.1 Percentage of external Center commitments achieved in the prior FY. **Goal—80% by due date.**
1.3a Number of new Positive Technology Transfers (PTTs) for R&T Base Program in the prior FY. **Goal**—60
1.3b Dollar value of PTTs for R&T Base Program in the prior FY
1.4 Number of citations in Science Citation Index for Science Programs or Projects in the prior FY: **Goal –1400**
1.5 Customer exit survey results on performance satisfaction for CSO4's major facilities during prior FY: **Goal—90% of highest possible.**
1.6 Percentage of new non-aerospace licenses for CSO4-developed technology during the prior FY which improve the quality of life in addition to having positive economic benefit for the U.S. (health, safety, environment). **Goal—25%**

Funder Value

2.1 Number of people presented information on CSO4 mission, capability, and contributions in the prior FY. **Goal—$400 mil.**
2.2 Dollar value of license revenues received by CSO4 in the prior FY. **Goal—$500K**
2.3 Major Projects led by CSO4 during the prior FY. **Goal – Two major projects, one significant element of a multi-year program and principal investigator for another program every three years.**
2.4 Annual Center funding in FY00 dollars. **Goal—$700M-$850M**

Organizational Value

3.1 Ratio of prior FY dollar value of R&T Base Program PTTs to end-of-year gross R&D funding guideline for those programs for that FY. **Goal—1.2**
3.2 Ratio of prior FY dollar value of R&T Base Program PTTs to total Full-Time Equivalents (FTE) charged to those programs for that FY. **Goal—$16K per FTE.**
3.3 CSO4 Science Program funding normalized by the corresponding CS workforce. **Goal—$500K/FTE.**
3.4 Occupational Injury lost time rates for the prior FY. **Goal – Less than 1%.**
3.5 Costing Performance for the prior FY. **Goal— 83% by EOY.**
3.6 CSO4 Employee Survey Results. **Goal—Survey every two years showing increased employee satisfaction.**
3.7 Annual employee and senior management scores on CSO4 Management Self Assessment. **Goal—Employee and Management self-assessment score within 5% of each other and +ve employee perception of management accomplishments.**

Figure 6-10: CSO4's Performance Scorecard

error that has unfolded over the several years since the scorecard's inception. For example, CSO4's management has experimented with several metrics and ways to measure employee productivity (e.g. positive technology transfer normalized with research dollars or FTE etc.) and continues to remain unsatisfied with its ability to get a good handle on the concept of individual productivity in a research setting. This, however, does not deter it from continually experimenting with different constructs and gradually evolving the scorecard into one that is a more accurate representative of organization's performance.

Performance monitoring is also carried out by using a time-series of performance data on each metric to learn about the efficiency or effectiveness of the underlying processes. This is formalized in CSO4's annual report that discusses each metric, its definition, rationale, and time-series of performance along with recommendations for possible corrective actions. The scorecard is also used as a means to assess on-going performance of the organization against pre-specified performance targets. The Figure 6-10 (above) presents the scorecard for the entire CSO4.

Each of the 33 sub-organizational units has its own scorecard that flows upwards into the center-level scorecard. These scorecards contain measures that are relevant to the sub-unit in question. For example, the performance scorecard of CSO4b—a basic research unit at CSO4—focused on academic citations/papers in its product user value category while that of CSO4a—a development unit at CSO4—focused on technology transfer.

CSO4's performance scorecards are designed to measure strategy—although, given the overall strategic inertia in the industry, they may measure organizational health more than they mirror the strategy. We found little churn in metrics because of changes in strategy. There is, however, some evidence that the scorecards do incorporate changes in strategy through the changes in initiatives that are linked to each of the metric on the scorecard. Another interesting observation is that performance measurement at CSO4 is not a continuous on-going exercise but rather a sporadic (often annual) ritual—a fact that might diminish its ability to influence individuals' performance on an on-going basis.

6.4.3—CSO4 compensates for its low incentive intensity by flexibility (through "menu of incentives")—CSO4 has an elaborate system—described by a set of published policies and procedure guidelines—of employee incentives, rewards, and recognition. While intrinsic factors like satisfaction with one's work is often considered the key motivator, several monetary and non-monetary awards, and recognition are available at the CSO4-level, parent agency-level, or at the level of the Federal government. Monetary awards include one-time bonuses and salary raises, group achievement awards, special rank and pay. Non-monetary awards include time-off awards, awards for specific accomplishments, quick appreciation awards, and medals of honor. The

majority of awards to individuals are made on the basis of subjective assessments of their performance that may or may not be backed by a quantitative measure. Managers, however, are assessed and rewarded—through a salary bonus of up to 25% of base pay—based on their unit's performance on the scorecard metrics. This ensures that there is adequate incentive for unit managers to pay attention to metrics on their scorecards.

In addition to the positive inducements, there is at least one factor that considerably dampens individuals' motivation to perform. CSO4, like other organizations constrained by the civil-service rules, also suffers from severe band compression around general scale-15 (GS-15) that results in dampened chances of promotion subsequently, thus considerably reducing a significant positive incentive for a large majority of senior scientists and engineers. This coupled with the inherent job stability as guaranteed by the civil service regulations significantly reduces the "carrot and stick" in CSO4's environment, namely, it fails to punish bad performance and does not provide strong enough opportunities for career progression and growth in recognition of good performance thus seriously affecting and undermining the morale and motivation of employees.

Whether or not this reward and recognition system net of the negative factors (described above) actually leads to improved performance on the job cannot be proved without a formal study. No such effort has been made at CSO4. The general impression that we got from CSO4's employees is that rewards may become important (ex-post) or at a significantly advanced stage into the project as they satisfy the individuals' need for recognition in the professional community. By providing a "menu of [such] incentives", CSO4 does a good job of compensating for its relatively weak pay-for-performance environment. A couple of words of caution are, however, in order. With so many awards to make, CSO4 must guard against the tendency of these becoming an "entitlement" rather than recognition of truly exceptional performance. Lack of fairness and transparency in the incentives structure can be a great de-motivator as is the presence of strong disincentives like the grade compression at the GS-15 level.

6.4.4—Although a Balanced Scorecard, CSO4's scorecard falls short on several structural features—On the face of it, CSO4's strategy-performance architecture that emphasizes three performance dimensions catering to its three key stakeholders maps onto the performance perspectives of the

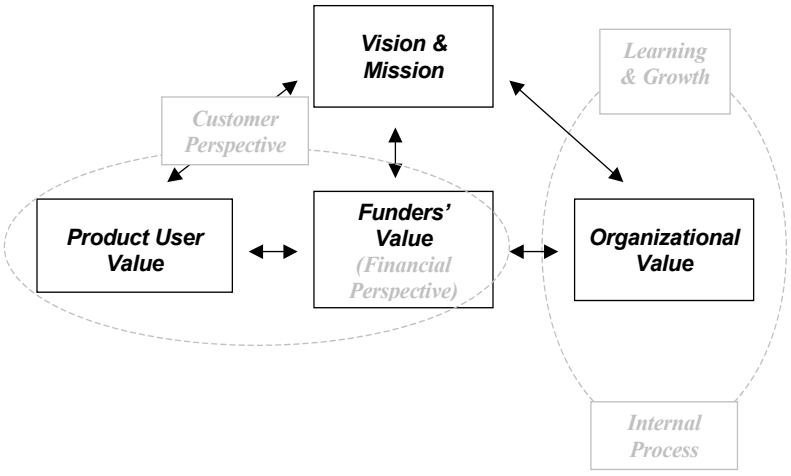


Figure 6-11: Balance in CSO4's Strategy-Performance Architecture

Balanced Scorecard. This mapping, however, lacks a one-to-one correspondence with the possibility that one or more aspects of performance might be compromised.

We also compared CSO4's performance scorecard with our own R&D Balanced Scorecard. The results are presented in Figure 6-11 The figure presents a picture of an organization that implicitly addresses all five performance dimensions even though it may not explicitly do so. The results on the simple strategy question suggest that CSO4's leaders and managers only believe three of the five key performance dimensions to be critical to its performance. It is no surprise that these three dimensions map one-to-one on the three critical success factors identified by CSO4.

However, a more in-depth look, using the detailed strategy questionnaire, presents a different picture. Here, we find that CSO4 pays at least as much emphasis to innovation management, and learning and management as it does on the other three dimensions. This strengthens and supports the notion that organizations, although they may not do so explicitly, implicitly take a multi-dimensional and balanced view of their performance.

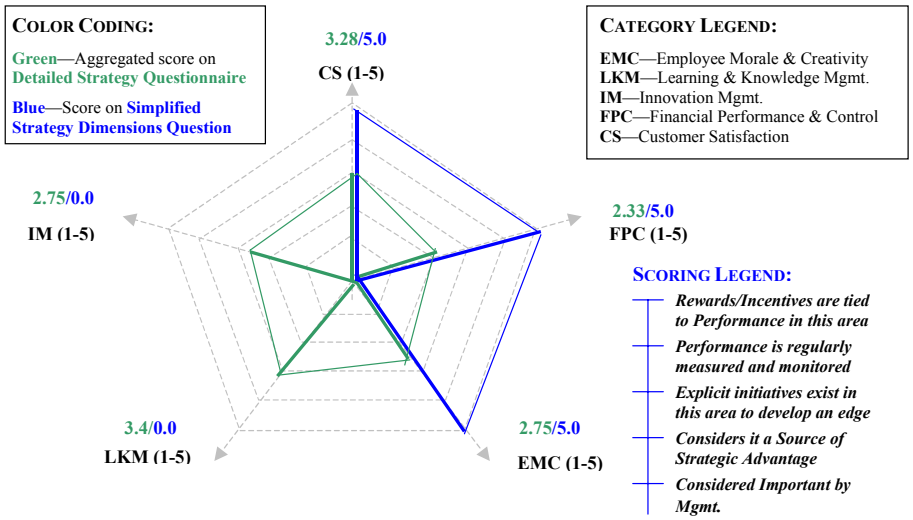


Figure 6-12: Strategy-Performance-Incentives Alignment at CSO4

CSO4's scorecard also lacks the explicit cause-and-effect linkage between various performance dimensions that is a hallmark of the Balanced Scorecard. While one may argue that the performance categories themselves can be interpreted to be causally linked, the scorecard itself does not emphasize this connection thus leaving it to the imagination of the employees to make that connection themselves. In addition, CSO4's scorecard's link with organizational strategy, on the one hand, and other managerial functions (e.g. budgeting, rewards/recognition etc.), on the other, is far from perfect. It also differs considerably from the traditional Balanced Scorecard in the manner it is used within the organization. CSO4's performance scorecard is often used as a part of its annual (or semi-annual) performance and strategy assessment exercise rather than a continuous tool for strategic planning and performance monitoring. This has the tendency of weakening its potential as

a motivator of performance that everyone within the organization would look at and benchmark his or her unit's performance against an on-going basis.

6.4.5—CSO4's Performance scorecard establishes the proof-of-concept for an R&D Balanced Scorecard in public sector—It also identifies several challenges that might be associated with such an undertaking—the most important of which is the need to integrate the Balanced Scorecard with the overall strategic management framework. Unlike the private sector where top-leadership of a company is autonomous (semi-autonomous) in charting the organization's strategy and future course of action, public sector managers are responsible to several stakeholders and are burdened by rigid civil service rules and procedural guidelines in their strategy-making efforts. We clearly see the effects of these limitations on the quality and extent of the implementation of the Balanced Scorecard at CSO4.

Another aspect on which CSO4's Balanced Scorecard performs poorly is its connection with the broader strategic management system. While the scorecard uses one of the several strategic frameworks employed at CSO4, it fails to link intricately with CSO4's real strategic planning to th extent that it comes across as an effort isolated from the rest of the organizational paraphernalia. Also, CSO4's strategic planning function is procedural, multi-faceted (in terms of sources of guidance) and repetitious (from year to year) to the extent of being bureaucratic—a fact that deprives it of the kind of nimbleness needed to implement "differentiated" strategies (a la Kaplan and Norton, 1996, 2001; and Porter, 1990) thus further limiting its ability to support a measurement framework designed to "measure the strategy".

CSO4's Scorecard also does not integrate as well with other aspects of CSO4's strategic management framework, namely, incentive systems, communications, budgeting, and performance improvement systems as a traditional Balanced Scorecard would have. Some of these shortcomings might be a reflection of incomplete implementation, though the scorecard has been in place at CSO4 for the last 6-8 years, or lack of understanding of the true potential of the Balanced Scorecard itself. Many, however, may just be a reflection of the inflexibilities of the public sector (R&D) environment. Fully exploiting its potential by integrating the Balanced Scorecard in all aspects of the strategic management systems, therefore, represents the true challenge for public sector organizations implementing a Balanced Scorecard.

CSO4's is a promising first-attempt at implementing a Balanced Scorecard within public sector R&D. Its leaders must, however, recognize that it has a long way to go before several of the hypothesized benefits could be derived from such an undertaking. CSO4 must build upon the early successes of (or lack of resistance to) its performance scorecard to deepen (in a qualitative sense) its

commitment to the Balanced Scorecard for it to find out whether this tool delivers the kind of performance improvements it claims to achieve in the private sector environments.

6.5—CASE STUDY #5:

BALANCED SCORECARD IN MULTI-SITE CENTRAL R&D ORGANIZATION OF A PRIVATE SECTOR FIRM

Case study organization # 5 (hereby, CSO5) is a multi-site central R&D organization of a major aerospace company. It is charged with developing advanced technologies, processes, and systems that could help create future aircraft and aerospace technologies that are safer, more capable and more reliable than today's - and can be designed, produced, and maintained in less time and for less money. CSO4 has a budget of about $1.5 billion with over 4500 employees spread across several R&D sites around the world. The case illustrates several important aspects of strategy-performance-incentives (SPI) systems at private sector R&D organizations that are of importance to our central thesis, specifically, the use of a Balanced Scorecard at such facilities. In addition:

- ***CSO5's has a differentiated strategic planning process that meets the specific needs of its various product categories.*** Several sources of guidance—sometimes disjoint with each other—impinge on the strategy-making process. CSO5, therefore, faces the challenge of putting in place a strategy that provides coherence to these multiple sources of guidance and makes it simple for employees to assess their own contribution to the organization's mission.
- ***CSO5 implements a Balanced Scorecard that emphasizes performance dimensions that only partially follow from the strategic planning process.*** The scorecard, however, does trigger a strategy dialogue within the organization that is considered a key benefit/feature of the scorecard development process itself. CSO5's performance architecture also includes a fairly extensive individual-level performance appraisal system whose link with the scorecard is, at best, weak and/or duplicative.
- ***CSO5 has an elaborate, high-intensity incentives environment aimed at rewarding individuals, financially and professionally, for their performance and making them a stakeholder in CSO5's success.*** While the (ex-ante) effect of incentives on actual performance is hypothesized, at best, they are considered necessary to retain talent within an organization—something CSO5 feels compelled to do--to match industry norms.
- ***CSO5's Balanced Scorecard, although balanced in several respects, does not incorporate several of the key structural features of the Balanced Scorecard.*** The most important of these include absence of a cause-and-effect model and its integration with the strategic management framework.

The distinguishing feature of CSO5's Balanced Scorecard implementation is its effort to "cherry pick" the features of the Balanced Scorecard that suits it circumstances the best—an approach that, if not done intelligently, can render the entire effort seriously compromised. The case illustrates a generic problem with implementation of managerial frameworks, like the Balanced Scorecard, in the private sector, namely, incomplete and faulty implementation almost always fails to achieve the desired results and is perceived as the failure of the management framework itself rather than lapses in implementation.

Following is a brief overview of CSO5's strategy-performance-incentives (SPI) system that illustrates its key features and lessons learnt.

6.5.1—Strategic planning processes at CSO5 caters for the differentiated needs of its diverse R&D portfolio—There are several sources of strategic guidance that impinge on the strategy-making process within CSO5. CSO5's Distinct Organizational Strategy (Figure 6-13, below) emphasizes making effective use of people and technology to create value for the parent. This value creation paradigm is reflected in CSO5's operating philosophy that consists of three distinct set of activities, namely, running healthy core businesses, leveraging strengths into products and services, and opening new frontiers—each of which depends on CSO5's ability to innovate and integrate.

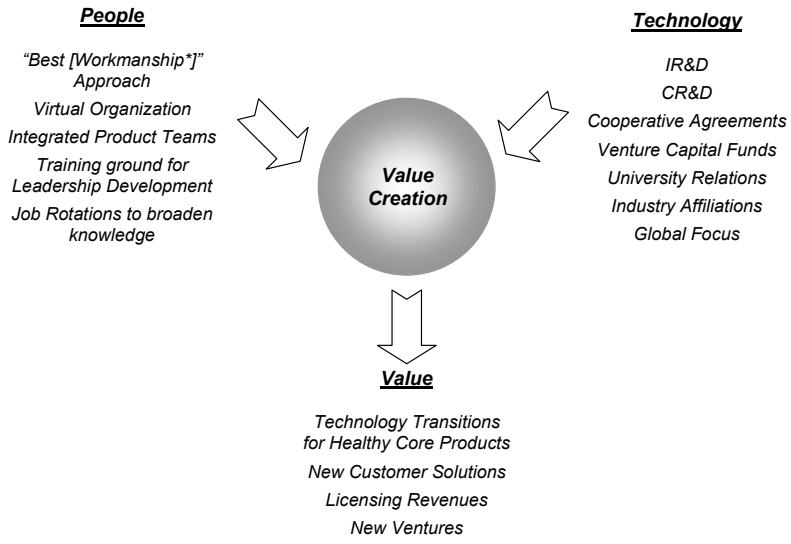


Figure 6-13: CSO5's Distinct Organizational Strategy

The strategy formulation process at CSO5 is varied, depending on the nature of the activity involved and aspirations and expectations of its diverse set of stakeholders. Significant differences may exist in the strategic planning paradigms of systems development, near and long-term technology development etc. For example, ***advanced systems development*** happens at a fairly advanced level of technology development that requires substantial commitment of resources. These are often make-or-break decisions for CSO5's parents, therefore requiring a corporate-wide decision. While research/project managers and individual scientists/engineers are involved in these decisions, their involvement is, at best, indirect and advisory. ***Near-term technology development*** decisions, on the other hand, require close cooperation between project managers and respective business unit managers. It is an iterative process that might originate at any one of the two poles but always ends

when a match between need and capability has been identified. The strategic planning process for ***the long-term technology development***, on the contrary, takes a mix of top-down and bottom-up approaches, and takes place at a much more inclusive (R&D Organization-wide) and participative environment with bench-level scientists and engineers involved in the strategic planning activity.

More generally, therefore, the involvement and influence of scientists and engineers over organizational strategy, and the autonomy and flexibility they enjoy in choosing what they work on, varies by project type. It is fairly extensive at the project level and tapers off as one moves to the more substantive systems-level decisions. At the level of the overall organizational strategy, for example, individual scientists and engineers enjoy little influence from year-to-year. They may, however, discuss and debate the organization's current strategic thrust, talk about it in all-hands meetings, and present ideas on new ways of doing things or doing new things altogether.

6.5.2—The Performance dimensions on CSO5's Balanced Scorecard only partially follow from its strategic planning process—CSO5's balanced scorecard, which rigorously follows a standardized pattern adopted corporate-wide, comprises three key performance dimensions, namely, growing the business, executing on existing business, and leveraging the best [of the corporate parent's capabilities]. The performance categories on CSO5's Balanced Scorecard as well as individual goals/metrics are picked, at least in principle, to reflect the organization's strategy. The strategic planning process, however, does not necessarily follow the tri-part characterization (discussed above). Although the basic structure of the scorecard does not change from year to year, one of the ways the scorecard reflects changes in

Category	#	Goal	Metric(s)	Weight (%)	YTD Performance	Year-end projection	Year-End Goal
1. Grow the Business	1.1	Technology Transition - Revenue	Increased revenue ($B)				
	1.2	Invention Disclosures	Year-End-Target: Total of 560 Disclosures				
	1.3	Orders/OTA captures - Performance to PW Op Plan	Orders and/or OTA captures ($M)				
	1.4	Capture key wins	Specific awards captured				
	1.5	Create/Exploit Opportunities	New contracts that were not in the Customer plans				
	1.6	Achieve [a mjor project's] SDD go ahead and successfully transition it to IDS	SDD Award and successful transition				
	1.7	Complete New Airplane technology and process development milestones in support of the go-ahead decision	Meet technical milestones				
Category	**#**	**Goal**	**Metric(s)**	**Weights (%)**	**YTD Performance**	**Year-end projection**	**Year-End Goal**
2. Execute Existing Business	2.1	Technology Transition - Cost	Cost reduction ($M)				
	2.2	CPAR Performance	Aggregate score on CPARs received				
	2.3	Customer Satisfaction Survey	Aggregate score on surveys received				
	2.4	Weighted SPI/CPI	Avg on major programs and thrusts				
	2.5	Performance to budget	Overhead and NBF budgets ($M)				
	2.6	Operating Earnings	Operating earnings ($M)				
Category	**#**	**Goal**	**Metric(s)**	**Weight (%)**	**YTD Performance**	**Year-end projection**	**Year-End Goal**
3. Leverage the best of CSO6	3.1	Leverage the best of Boeing	Improvement in ESI				
	3.2	Development of people	10% people rotation				
	3.3	Develop effective leaders	% of Mgt Opportunities filled by Diversity and Succession plan candidates				
	3.4	Support to Business Units	Aggregate score on surveys received				
	3.5	Communication	Improvement in Survey Communication score				
				Score			

Figure 6-14: CSO5's Balanced Scorecard

strategy is the use of metrics specific to research and development projects that reflect the organization's current strategic thrusts. Figure 6-14 presents CSO5's performance scorecard. Color codes (green, blue, yellow) reflect the level of progress made against pre-specified performance goals.

CSO5 employs a cascaded structure with each level of scorecard dovetailing into the scorecards of the higher organizational levels. The process begins by developing the top-level scorecard for the entire organization. Once the top-level scorecard has been developed, each of the Directors of the various sub-organizational units are asked to develop their own set of goals and scorecards reflecting their contribution to CSO5's commitments to the corporate parent. This process trickles down, one organizational layer at a time, with Directors taking their scorecards and that of their bosses to managers who in turn take theirs to their own immediate subordinates. This process of cascading sets into motion a "strategic dialogue" that is considered important by the top-leadership of CSO5.

The scorecards thus created also become the center point of organization-wide conversations about performance. Every month each one of the metrics on the scorecards is reported, year-to-date performance is reviewed, and a year-end projection is made. Also, issues that might be causing aberrations in performance are identified and actions needed to address those are prescribed. At Quarterly Business Review meetings as well, CSO5's top management (i.e. president and his direct reports) review the organization-level performance scorecard. Finally, scorecards are also displayed at all-hands employee meetings to let people know of where CSO5 stands vis-à-vis its goals and objectives. The scorecards at CSO5, thus form the basis of a fairly extensive performance measurement architecture that is widely used and referred to within the organization.

6.5.3—High-intensity incentives system rewards individual performance designed to make them stakeholders in CSO5's success—CSO5 has a fairly elaborate incentives and rewards system comprising non-financial/professional and financial rewards to recognize performance. The purpose, intent, size, eligibility and other important dimensions of each award (or awards scheme) are clearly laid out in the corporate policy manuals. Financial rewards include cash awards, special incentive awards, and various forms of stock options and employee ownership plans. Non-financial awards include team awards and corporate pride awards designed to acknowledge the contribution of teams and individuals to the organizational goals.

The general feeling at CSO5 seemed to be that while rewards may not be the primary motivator, they do have an influence on the scientific staff. One of the senior managers that we spoke to highlighted that scientists and engineers that work at CSO5 are a very aware lot, from a financial standpoint, and track how they are paid vis-à-vis their peers elsewhere. The manager stressed that

many would actively complain if they believed they are being paid less than their opportunity cost. This individual was of the view that ultimately the incentives "alone" may not be a powerful enough motivator but "the entire package—the environment, the challenge, the compensation, the rewards/recognition, and the mere intrinsic satisfaction of doing what is useful and what one likes doing—is key to performance in the workplace."

CSO5's leadership and management responds by providing a "menu of incentives" and rewards that tends to influence employees performance (ex-post) by creating a workforce that is motivated, feels appropriately appreciated and acknowledged, and well taken-care-of. This positive feeling, the management believed, is likely to have a definite influence on individuals' performance.

6.5.4—CSO5's Balanced Scorecard falls short of fully adopting some key structural features— While CSO5's Balance Scorecard does take a multi-dimensional view of organizational performance, it certainly falls short of balance between key performance dimensions. Specifically, while the three-part categorization might include certain elements of the performance perspectives of the Balanced Scorecard, it fails to clearly account for at least three of the four perspectives identified in the Balanced Scorecard literature, namely, internal process, customer satisfaction, and financial performance. CSO5's existing performance architecture does cater to the time-dimension through its emphasis on growing current vs. future businesses. It, however, fails to account for balance across the input-process-output-outcome continuum.

We asked CSO5's leaders and managers about the performance dimensions deemed critical to performance (explicitly) and tried to assess, through a detailed strategy questionnaire, if they implicitly did so. The results are presented in Figure 6-16 (below). They represent an organization that is considerably well-balanced along the five hypothesized performance dimensions, even though the latter do not explicitly feature as performance categories (a la the Balanced Scorecard).

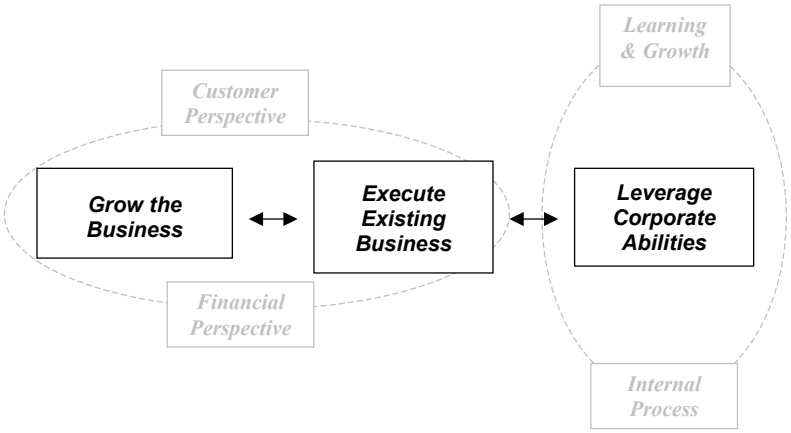


Figure 6-15: Balance in CSO5's Strategy-Performance Architecture

In terms of the structure and specifics of implementation, CSO5 presents a mixed picture. While its performance scorecard does exhibit some characteristics of the traditional Balanced Scorecard (e.g. balance between dimensions of performance, and cascading throughout the organization), it falls short on several others (e.g. cause-and-effect logic, linkage with overall strategic management system.) Under these circumstances, the ability of CSO5's performance scorecard to act as tool for strategic learning is severely limited.

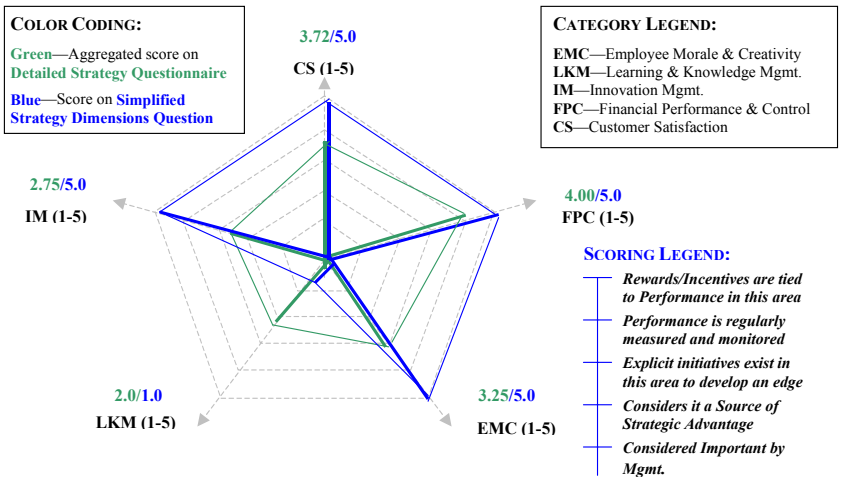


Figure 6-16: Strategy-Performance-Incentives Alignment at CSO5

The most glaring omission is the lack of a clear linkage between CSO5's performance scorecard and the overall strategic management framework (e.g. budgeting and performance improvement efforts) or even the strategy itself. We did not find any evidence to suggest that each of the three differentiated strategic planning modes used the three performance dimensions of the performance scorecard to guide the strategy-development process. In all likelihood, CSO5 uses the strategy-performance systems as two relatively disjoint managerial frameworks rather than a single integrated whole whose various components speak to each other. Therein lie the most significant opportunities for CSO5 to improve the effectiveness of its strategy-performance-incentives (SPI) systems.

6.5.5—"Cherry Picking" certain aspects of the Balanced Scorecard without giving due thought to its impact on completeness of the larger framework can render it severely compromised—This case illustrates several important aspects of the Balanced Scorecard as applied to an R&D organization. For a start, CSO5 is a private sector R&D organization which relieves it of several of the inflexibilities (e.g. civil service regulations, bureaucratic organizational processes, multiple stakeholders etc.) encountered in public sector labs that might hinder complete implementation of the Balanced Scorecard. As a result, therefore, we see a more streamlined strategy-making process and a better connect between individual performance and incentives. But it also illustrates several other features of R&D labs that are common across both public and private sectors that may still have to be resolved before the Balanced Scorecard may be implemented completely and comprehensively. These include, among others, the lack of a clear bottom-line, multiple criteria of assessing R&D performance, and the predominance of intrinsic rather than extrinsic motivators.

Most importantly, however, this case illustrates an important point that has been repeatedly emphasized in the management literature. This refers to the need for implementing all essential structural features of the management/measurement framework in question instead of cherry-picking the convenient ones. CSO5, like any other typical organization, only implements certain aspects of the Balanced Scorecard and not others. For example, the strategy-making process at CSO5 is still very compartmentalized and isolated from the performance measurement processes. This leaves open the possibility that CSO5 still conducts strategy-formulation and performance measurement as two separate activities rather than a part of a single overall strategic management process. Cherry-picking the manner and extent to which different structural features of the Balanced Scorecard get implemented at CSO5 can have serious bearing on the sort of results that would be achieved from such an implementation.

CSO5, despite being several years into an implementation of the Balanced Scorecard, has not seen the sort of dramatic performance breakthroughs hypothesized by the initiators of the Balanced Scorecard. This coupled with the fact that the implementation itself is somewhat patchy serves to illustrate an important methodological issue. Among these is the tendency to attribute the inability to achieve hypothesized results to the failure of management framework itself rather than the failure to implement the management framework in its true shape and spirit. CSO5's example illustrates the need to think about the "completeness" of the overall framework, and accordingly adjust ones expectations of results.

6.6—CASE STUDY #6:

STRATEGY-PERFORMANCE ARCHITECTURE AT A UNIVERSITY DEPARTMENT & RESEARCH CENTER

Case study organization # 6 (hereby, CSO6) is an institution of higher learning with revenues of around $150 million and R&D expenditure of around $30 million—30% of which is spent on basic research, 20% in applied research, and the rest on product/process improvement, new product development, and test and engineering. We looked at two strategy-performance-incentives systems of two sub-units of CSO6, namely, an academic department [of mechanical engineering] (hereby referred to as, CSO6a) and an affiliated research center of similar technical specialty (hereby referred to as, CSO6b). The case illustrates the differences in strategy-performance-incentives systems that might arise due to the differing requirements and nature of work at these two sub-units, even though the overall strategic and incentives environment is essentially the same. CSO6 was also unique in that, contrary to other case study organizations, its leadership had felt a need to bring subunits, CSO6a and CSO6b, through better alignment of their strategy and performance systems and was in the midst of major re-thinking in that regard. Towards that end, CSO6 enabled us to

study an SPI system that was in the state of being redefined. Specifically, this case displayed the following characteristics:

- ***CSO6's differentiated strategic planning processes, reflecting the differing requirements and traditions of the two sub-units, run the risk of achieving local maxima at the expense of cross unit synergies.*** CSO6 faces the challenge of streamlining the strategic planning processes and activities of the academic department and the affiliated research center in order to invite cross germination of staff and ideas, explore synergies across the two units, and avoid duplication of effort.
- ***CSO6's performance measurement system comprises traditional peer review at the academic unit and a variant of MBO at the research center.*** The fact that both these systems emphasize aspects of performance that can be different from (and at times even counterproductive to) each other is a major challenge for the designers of an integrated performance architecture that combined these metrics and allows CSO6's leaders to make trade-offs.
- ***CSO6 incentives structures are fairly low-powered and traditional with greater emphasis on intrinsic sources of motivation rather than on extrinsic awards/rewards.*** Promotion and award of tenure are the greatest motivators that force individuals to perform as per pre-set performance expectations.
- ***CSO6 does not implement a Balanced Scorecard but its strategy-performance systems implicitly aspire for balance across performance dimensions.*** The most important of these include absence of an explicit cause-and-effect model and the lack of balance along several organizational dimensions.

The distinguishing feature of CSO6 Strategy-Performance-Incentives (SPI) systems is its aspiration to cater to the differing needs of various sub-organizational units (CSO6a, CSO6b). The case illustrates the problems associated with trying to manage diverse (in terms of nature of work, and performance expectations etc.) sub-organizational units. It emphasizes the need to use an interlinked (or uniform) strategy-performance-incentives architecture to explore synergies across sub-organizational units.

Following is a brief overview of CSO6's strategy-performance-incentives (SPI) system that illustrates its key features and lessons learnt.

6.4.1—A differentiated strategic planning process, reflecting the differing requirements and traditions of the two sub-units, may achieve local maxima at the expense of cross unit synergies— The strategic planning process at CSO6 is a mix of centralized (university-wide) guidance and decentralized (department-level) planning. Several sources of strategic guidance (e.g. a mission, a vision, a set of guiding principles, and a long-range strategic plan) impinge upon the strategic planning process as it unfolds at various levels. The process, featuring around collegial consensus

building and long-established traditions of academic disciplines in question, also reflects a fair degree of permanence and stability. At the organizational level, it hinges upon a traditional long-range planning and implementation plan termed "Action Agenda for 2010" which was developed and approved by its Board of Control and is drawn up through a series of consultative meetings. An Action Plan supporting the agenda is also developed to identify activities and goals to be achieved in short-, mid-, and long-term. This action plan is responsible for tying the institutional mission and vision with its budgetary processes. This annual strategic planning process is also seen as an opportunity to look at the strategic portfolios, review goals, objectives, and strategies, and develop methods of measuring progress toward achieving the unit's strategic goals.

At the level of an academic department, for example, CSO6, the process is carried out in a much more collegial and tradition-rich manner. Each department periodically holds strategic planning meetings (at least annually) that go through a process of stocktaking during which its current progress, goals, and future direction are discussed. For example, during one such strategic planning session, it may transpire that the department's performance is below expectations in terms of research output. Reasons for the lagging performance may be discussed and appropriate measures (e.g. hiring of younger faculty, creation of dedicated seed funds for research etc.) may be discussed and adopted at such meetings. More recently, however, these meetings have also included discussions on department's obligations towards and performance on institutional goals.

This process is, in principle, quite participative whereby the entire faculty of the department is allowed to deliberate on the issues at hand. In practice, however, it is much less than that. Several faculty members are not interested in participating in the process. Several others may feel that they have little say in the process, given the involvement of department's heavy weights and the long-honored traditions of the department and the discipline. Still others may be very interested and involved in the details of the process and feel greater ownership of the process as well as the objectives derived from such a process.

At dedicated research center (e.g. CSO6b), on the other hand, the strategic planning process follows a much more hierarchical and corporate-like approach whereby the center director reviews, with his or her closest associates (or direct reports) or in a more open and participative manner, the past performance of the center and, given the need, come up with ambitious plans to address the shortcomings. At CSO6b, for example, the strategic planning process takes place in a much less participative fashion with the center director and an administrative committee of five-odd individuals taking stock of the situation and making key decisions vis-à-vis the future of the research center.

Although being a part of the same overall organization (i.e. CSO6), there are several overlaps and contradictions between the strategic planning processes at these two sub-units. Developing a well-

respected research capability, for example, is an organizational priority for both CSO6a and CSO6b. Publishing research, however, may only be a top-priority for the academic unit and may play second fiddle to the more important need of bringing in research dollars at the dedicated research center. To the extent that a well-respected research capability can do both i.e. enable researchers to publish and bring in more research dollars, it can be a source of synergy between the two sub-units. This, however, is not always the case. Writing and publishing papers can divert ones attention from seeking project money or finishing existing projects, especially if the latter are downstream of the R&D spectrum (i.e. development and testing etc.)—an area that is the largest source of R&D dollars for CSO6b. This calls for a unified rather than differentiated strategic planning process that is designed to aspire for global rather than local maxima.

6.6.2—Performance measurement system at CSO6's sub-units are challenged by the differing (at times, even counterproductive) requirements of the two sub-units—CSO6 has a multi-tiered architecture comprising three layers of metrics, namely, thematic/national measures, goal measures, and unit measures. National or thematic measures are those that allow CSO6 to compare its performance against other peer institutions. They are "top level" and cross-cutting and may measure more than one CSO6 strategic goal. The goal measures are university level measurables and are purposely few in number. They are holistic, university-wide, appropriate for Board or university executive use, small in number, and developed once per strategic plan. They are rarely, if ever, revised. The unit-level measures are developed and used by individual units to assess unit progress toward unit strategic goals stated and described in the Unit Strategic Portfolios. Taken together, these metrics are routinely collected and used for different purposes, namely, early-warning, calibration and benchmarking etc. These include input (e.g. expenditure, $-awards), output (e.g. publications, students, licensing activity etc.) and process (e.g. proposals submitted etc.) measures. These measures are collected at multiple (at least three) levels, namely, university, college, and academic units and are linked through the process of cascading.

The performance measurement system at the academic unit level (i.e. at CSO6a) is the traditional academic peer review/tenure track system. Individuals are expected to do research and publish papers, teach a specified number of courses, and get involved in service to the institution and the profession. Their performance is measured accordingly. The process is qualitative and highly collegial. Some objective measures of performance, however, are also defined. The expectations of performance are clearly defined, right from the junior most (pre-tenure) to the senior most (professors) level. A new assistant professor, for example, is expected to bring in $100-150K/year worth of research support, graduate PhD students, and write 1-2 journal articles/year before he or she could become eligible for promotion to the next level.

At the affiliated research center (i.e. CSO6b), on the other hand, more objective criteria of performance are used. Measures such as dollar value of research contracts, number of students employed, number of customer visits, and number papers and presentations are considered most critical measures of performance. The key emphasis here, however, is to keep everyone at the center employed and to ensure the stability of funding. Although there are no hard and fast rules or targets, every senior researcher is expected to bring in 1-2 times of his or her salary towards the common pool of resources that are used to run projects at CSO6b. These quantitative measures, however, are used in a flexible manner. Year-to-year variations in publication and research sponsorship are not only common but also considered a natural phenomenon. Performance measures are often averaged over several years (3-6 year averages are a norm in the tenure process) to account for these year-to-year variations.

6.6.3—CSO6 incentives structures are fairly low-powered with greater emphasis on intrinsic sources of motivation rather than on extrinsic awards/rewards—The incentives structure at CSO6 is typical of one found at an academic institution. Salary raises are fairly straightforward, and are generally based on individuals' merit. A merit pool is established each year depending on the institution's financial health and ability to pay a raise. For example, the merit pool for the preceding year was at 2.9% average salary increase level. It was 0% and 1.5% in the years preceding the previous year. By any measure, however, promotion to the next level (at both CSO6a and CSO6b) and granting of tenure (at CSO6a) were the strongest performance incentives.

Financial incentives, except for the ones cloaked as professional incentives (discussed above), are somewhat rare and limited by state funding statutes and restrictions imposed by funding agencies. One exception is the IP royalty-sharing arrangement that allows researchers to claim up-to a third of the gross royalty received by the university from their patents. CSO6 uses other informal means to reward and encourage better performance from individual researchers. For example, for younger faculty, travel money to attend and present at conferences and professional meetings is considered to be a strong motivator. So are start-up funds and departmental support (e.g. relieving a faculty members' teaching burden) to write research proposals.

CSO6's leadership believes that while financial rewards do seem to matter generally, their impact is fairly limited in an academic setting. For example, it may not be possible to use rewards and recognition to double or triple an individual's intellectual productivity. Freedom and flexibility, intellectual autonomy, and respect of peers are considered the greatest motivators of performance. Money can be a strong motivator or not at all. For example, there was a sense among the researchers that money in the form of additional research dollars can always be a useful motivator of performance but money in ones pocket may or may not be one.

6.6.4—CSO6 does not implement a Balanced Scorecard but its strategy-performance systems implicitly aspire for balance across performance dimensions— While CSO6's strategy-performance architecture does acknowledge the notion of performance multi-dimensionality it does not incorporate, equally, the four key performance perspectives of the Balanced Scorecard. Using CSO6's seven goals as proxies for performance dimensions, we find a less-than-perfect correspondence with the BSC performance perspectives. While this presents (Figure 6-17) the picture of an organization that takes a "balanced" view of performance, it is in-fact only designed to achieve balance across different dimensions of output performance thus leaving behind other aspects of balance, namely, input, process, output, and outcome measures as well as leading and lagging measures.

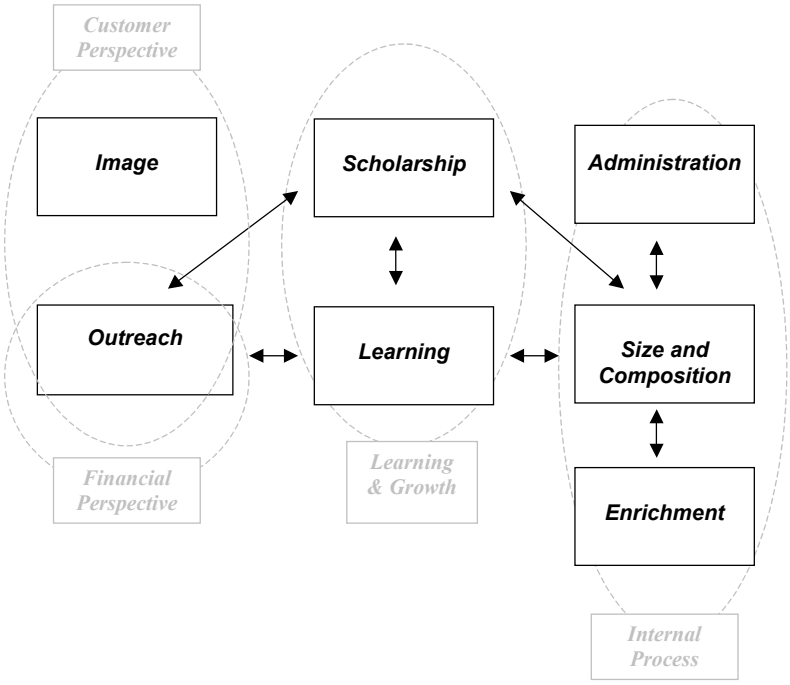


Figure 6-17: CSO6's Strategy-Performance Architecture

Figure 6-18 presents another view of balance in CSO6's strategy-performance-measurement systems. It maps CSO6's strategy-performance architecture on our proposed R&D Balanced Scorecard. It presents the view of an organization that does aspire for balance across multiple dimensions (perspectives, in Balanced Scorecard terminology) of performance. Some of the dimensions of performance (e.g. employee morale and customer satisfaction) are not measured directly (e.g. through a customer satisfaction survey) but are measured indirectly through other constructs (e.g. retention rates, graduation rates etc.) The detailed strategy questionnaire that looks at sub-dimensions within each of the five major performance dimensions also validates the findings of the simplified strategy questionnaire—thus bringing more credence to the results.

CSO6's strategy-performance architecture does not incorporate some of the structural features of the Balanced Scorecard (e.g. an explicit cause-and-effect performance model and strategy-mapping) and incorporates others only partially (e.g. balance across performance dimensions, cascading) – thus limiting its ability to influence individual and organizational performance in ways that are characteristic of an organization implementing a Balanced Scorecard.

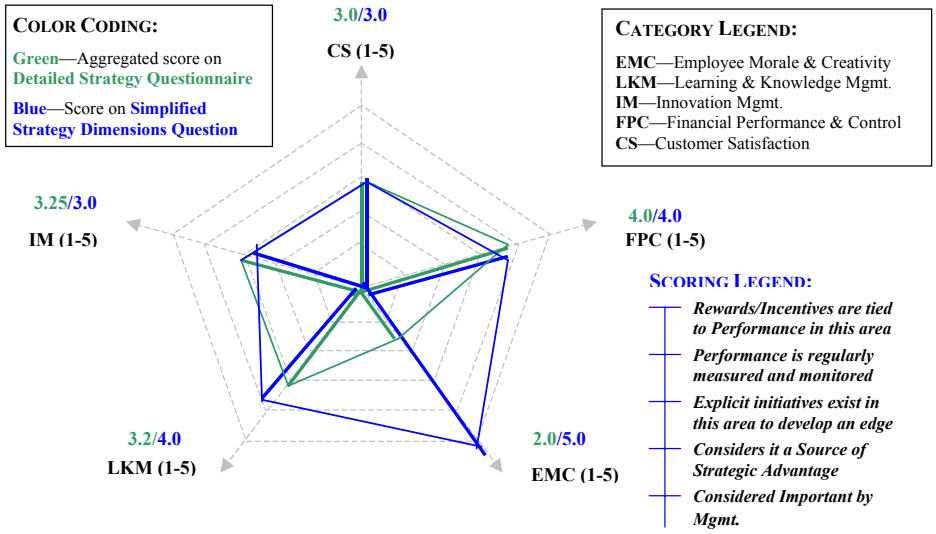


Figure 6-18: Strategy-Performance-Incentives Alignment at CSO6

6.6.5—The distinguishing feature of CSO6 Strategy-Performance-Incentives (SPI) systems is its aspiration to cater to the differing needs of—with the resultant lack of uniformity across—various sub-organizational units—While both CSO6a and CSO6b are sub-units of the same organization (CSO6) and hence subject to same organizational factors (e.g. strategy-making processes, incentives systems, service regulations etc.), they have developed their own strategy-performance architectures to meet their unique needs and performance expectations. This case study was conducted against the backdrop of a need to enable the two sub-units to work together and capitalize on synergies between them. The case clearly illustrates some of the key problems that arise due to differentiated strategy-performance architectures within the same organizational environment. CSO6a—an academic research unit—is driven, primarily, by the goal to do research that pushes the frontier of knowledge in its discipline. CSO6b—a soft-money research center—is driven, primarily, by the goal to attract research funding (generally, downstream of the R&D continuum e.g. for development and testing projects). The two sub-units are driven by different strategy-formulation, performance measurement, and incentives structures that meet their unique requirements respectively.

The problem arises when these different strategy-performance architectures optimize performance at the sub-unit level (i.e. local optima) but not at the overall organization-level (i.e. global optima). For example, the researchers working at the academic department (CSO6a) may be served well by a

strategy-making process, performance measurement, and incentives system (academic peer review and tenure process etc.) that forces them to do research at the cutting edge but discourages them from participating in the activities of the research center (CSO6b) that involves little cutting-edge research and opportunities for publication. Conversely, the researchers at the research center (CSO6b) may be served well by a corporate-like strategy-performance-incentives architecture that forces them to seek R&D dollars but leaves them little time and opportunity to write publishable papers. These problems of differentiated strategy-performance-incentives architectures may be addressed through a uniform strategy-performance-incentives system at the level of the overall organization.

A Balanced Scorecard-type cause-and-effect model of performance that clarifies the relationships between various types of outputs (e.g. papers vs. research dollars, research papers written at CSO6a vs. those written at CSO6b etc.) and their relationship to the overall organization may help in determining the right balance between the various kinds of outputs. Performance expectations and incentives systems may then be adjusted to reflect that balance. The resultant hybrid strategy-performance-incentives architecture must arise from CSO6's overall strategy and form the basis of key aspects of its strategic management system. A Balanced Scorecard-type system thus conceived has the potential to capitalize on synergies between various sub-organizational units by making explicit the cause-and-effect relationships, incentivizing performance by improving the strategic focus, and encouraging right behaviors by individuals.

In the chapter that follows, we look across the six cases in an attempt to present evidence in support of or against the hypotheses identified earlier and draw generalizable conclusions about the strategy-performance architectures in R&D organizations, both Balanced Scorecard-based and otherwise.

CHAPTER # 7

UNDERSTATING STRATEGY-PERFORMANCE-INCENTIVES (SPI) SYSTEMS IN R&D: A CROSS CASE COMPARISON

In the previous chapter, we described the salient features of the six organizations studied as a part of our case study design. These descriptions were set within the contexts of each of these six organizations and focused on key constructs of importance to this research, namely, strategy-formulation, performance measurement, and incentives systems and their inter-linkage with each other. In this chapter we look at the cumulative evidence that can be derived from these six organizational cases in a manner that would allow us to draw some generalizations across them. W also—for the purpose of corroboration and/or further enrichment of the theory—supplement the case study evidence with interviews with R&D Directors and research managers from 32 organizations across a wide spectrum of organizational types, R&D activities, and mission requirements. Together, the case study evidence combined with insights from interviews of a relatively wider set of organizations has the potential to enhance our understanding of the strategy-performance-incentives systems and their various contexts and contingencies within R&D organizations.

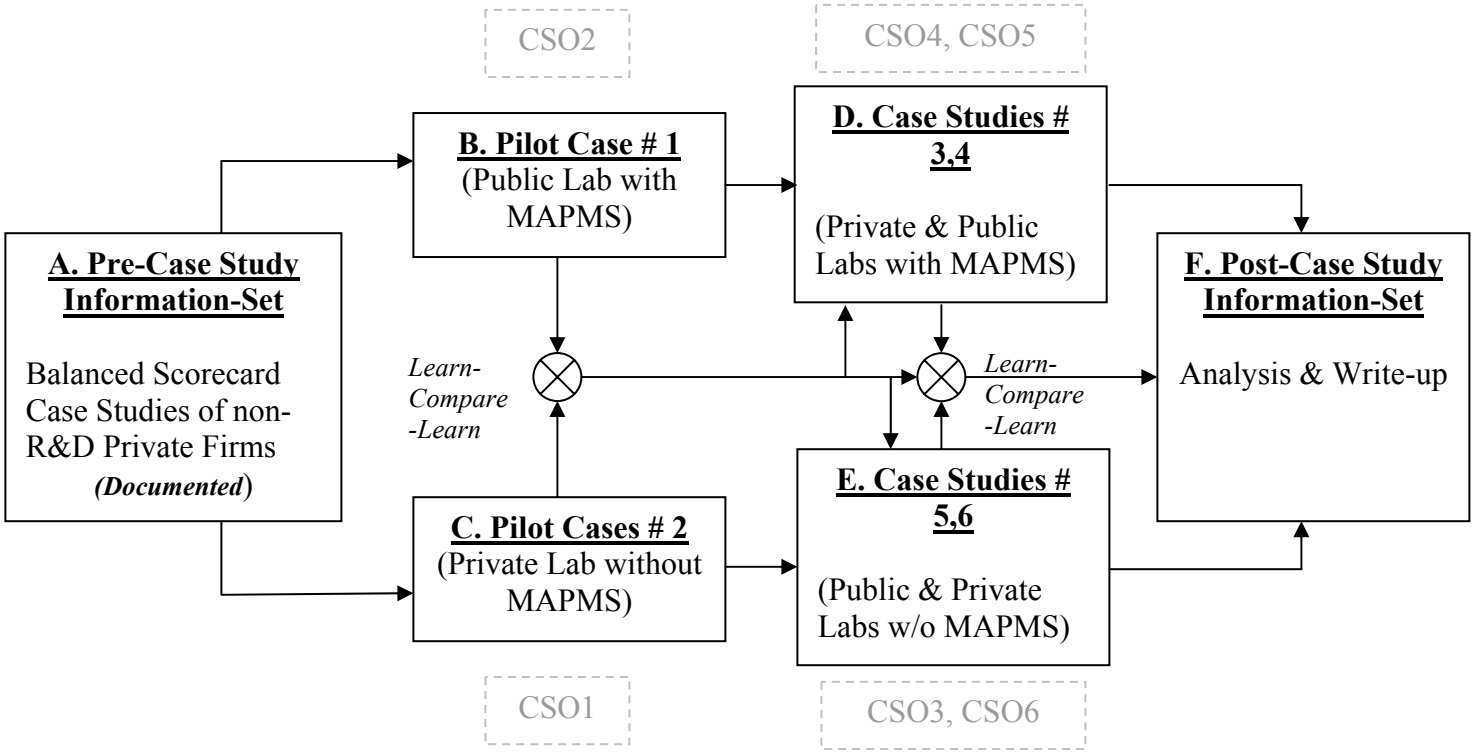
7.1—REVISITING THE CASE STUDY DESIGN

Before we begin our discussion of the cumulative case study evidence, it is important to revisit aspects of the case study design itself that we draw upon in this analysis[83]. We capitalize on the organizational diversity within a multiple case study design framework to not only cover a cross section of R&D organizations and activities but also to test various theoretical propositions. The case study design is recreated in Figure 7-1 (below) from chapter-4. The idea underlying the design is to be able to control qualitatively for certain variables as we try to learn and make generalizations about the constructs of interest to us.

In the context of the current investigation, the case study design controls for a key characteristic of these systems namely, their multi-attribute character while varying two other relevant dimensions, namely, the organizational structure (i.e. public vs. private vs. academic) and the type of R&D performed (i.e. basic/applied vs. development). This analytical strategy is akin to intelligent experimental design in scientific research and has the potential of allowing us to exploit the diversity of the organizational space to learn about a broader range of phenomena than would have

[83] This discussion would largely draw upon the discussion in 4.2.4.

been possible had such an analytical design not been adopted. The accompanying table in Figure 7-1 briefly illustrates the types of cross comparisons that are made possible through the use of the adopted research design. The figure also identifies (in dotted gray boxes) specific case study organizations that make up the various components of the overall information set (namely, B, C, D and E).



COMPARISONS	KEY ANALYTICAL OBJECTIVES ACHIEVED
A-B	Sheds light on differences between MAPMS[84] in business vs. R&D
A-C	Sheds light on why R&D might be different from normal business firms?
C-B	Sheds light on how R&D firms using MAPMS are different from those that don't?
B-D	Builds upon findings of the pilot in B. (achieves literal replication)
C-E	Builds upon findings of the pilot in C. (achieves theoretical replication)
D-E	Builds upon findings of C-B
D-E-F	Tests hypothesis at the cross case level.

FIGURE 7-1: THE PROPOSED STUDY DESIGN & ANALYTIC PLAN

The second important feature of the proposed case study research design was the use of embedded units of analysis. The embedded units of analysis allow us to further control for a host of variables

[84] Stands for multi-attribute performance measurement systems.

not possible, normally, in the case study approach. While specifying a unit of analysis—in this case, the strategy-performance-incentives systems within case study organizations—requires a researcher to limit his or her analysis and data collection activities within predefined boundaries in a multiple case study context, the embedded units of analysis allow us to further focus on and conduct a more detailed investigation of a specific construct/issue of interest within the larger unit of analysis (Yin, 1984, 1993).

In the present context, the embedded unit of analysis is a specific phenomenon that could be observed in several of our case studies, namely, the differential impact of strategy-performance-incentives (SPI) systems on sub-organizational entities that may differ with reference to the type of R&D performed. More specifically, while we focused, more generally, on strategy-performance-incentives systems within our case study subjects, we also looked for opportunities to see how the implementation of such systems varied across sub-organizational units within a particular case study organization. In four out of six case study organizations (i.e. CSO2, CSO3, CSO4, and CSO6) we found such sub-organizational units (e.g. a technology lab affiliated with a research lab, an affiliated research center associated with a university department, a basic research center within a larger development organization, and an administrative department subject to same SPI systems as an academic/research department within a university). This is akin to creating *"cases within a case study"* thus providing us with greater flexibility, range of phenomena to study, and greater possibility of generalizing through literal or theoretical replication[85].

Whenever we found such sub-organizational units, we capitalized on the opportunity afforded to us of controlling for virtually every organizational-level construct and observing the differential impact of the same set of strategy-performance-incentives system characteristics on two distinct entities that differed only in terms of their mission and/or the type of R&D activity performed. For example, while our broader case study design allowed us to observe how the implementation of similar types of strategy-performance-incentives systems (e.g. multi-attribute vs. non-multi-attribute) varied across different organizational contexts that were either engaged in research or development, the embedded unit of analysis allowed us to observe how the implementation of the same strategy-performance-incentives systems differed across sub-organizations within the same organizational context. This ability is important to further fine-tune the discussion on the effect of contingency

[85] Yin (1984, 1993) describes in some detail the concept of generalization in a case study format. According to this view, the predominant analytic strategy to achieve generalization in the case study context rests on the principle of replication rather than a sampling logic. While the sampling logic depends upon the generalization of a phenomenon observed in a sample to a pre-defined population based on principles of statistics, the replication logic is analogous to that used in multiple experiments. The replication logic draws heavily on the underlying theory. In the replication logic, each case studied—as it unveils—must either a) predict similar results (a literal replication) or b) contrasting results but for predictable reasons (a theoretical replication) vis-à-vis the predetermined theory.

factors (e.g. organization size, type of research, mission etc.) on the choice and implementation of strategy-performance-incentives systems.

7.2—Salient Features of Case Study Organizations (CSOs)

The selection of case study subjects for this research was based on several criteria, the most important of which being compliance with the requirements of the case study design (Figure 7-I, above), willingness on the part of prospective subjects to provide desired level of access to the organization, and representation of diversity within the target population. The selection was done through the help of a detailed survey instrument and an hour-long pre-screening interview. We selected six organizations to become subjects of the case study phase of our research. A brief snapshot of the case study subjects (numbered as CSO1 to CSO6) is given in Table 7-1 (below).

Table 7-1 – A Brief Snapshot of Case Study Organizations (CSOs)

CSO#	Org. Type	Sector	Type of R&D[86]	Embedded Unit[87]	PMS System[88]
CSO1	*Private*	*Medical*	*AR/TD*	*No*	*SS, SG, MBO*
CSO2	*Public*	*Academic*	*BR/AR*	*Yes*	*BSC, PR*
CSO3	*Private*	*Info Tech.*	*BR/AR/TD*	*Yes*	*MBO, PR*
CSO4	*Public*	*Aerospace*	*BR/AR/TD*	*Yes*	*BSC*
CSO5	*Private*	*Aerospace*	*BR/AR/TD*	*No*	*BSC, MBO*
CSO6	*Public*	*Academic*	*BR/AR*	*Yes*	*MBO, PR*

Table 7-1 indicates the diversity of our case study subjects. These comprised organizations engaged in R&D spread across the entire spectrum of research and development activities. Our case study

[86] For type of R&D work, BR represents Basic Research, AR represents Applied Research, and TD represents Technology Development.

[87] An embedded unit of analysis represents a situation within some case study subjects where differential impact of an intervention may be observed on a sub-unit or a sub-phenomenon within the broader (organizational) context.

[88] For Performance Measurement System (PMS), SS represents Six Sigma, SG represents Stage-gate process, MBO represents Management-by-Objectives, BSC represents Balanced Scorecard, and PR represents Peer Review.

TABLE 7-2– SALIENT ORGANIZATIONAL FEATURES OF CASE STUDY ORGANIZATIONS (CSOS)							
CSO#	**ANNUAL R&D EXPENDITURE**	**SIZE [EMPLOYEES]**	**ORG. ENTITY STUDIED % Ph.Ds/MDs/M.Phils**	**R&D MIX [BR:AR:TD]**	**PERFORMANCE RATING [SELF]**	**EMBEDDED UNIT**	**STRATEGIC POSTURE**
CSO1	*$10mn.*	*<20*	*Overall Co. R&D 20%*	*1:5:94%*	*Above Average*	*No*	*Tech. Excellence Ops. Excellence*
CSO2	*$125mn.*	*>1000*	*Overall Co. R&D 40%*	*60:18:22%*	*Above Average*	*Admin Dept. Acad. Unit*	*Tech. Excellence Customer Responsive-ness*
CSO3	*$8mn.*	*20-100*	*Division Lab (R-Lab) 85%*	*75:25%*	*Top-Quartile*	*Research Lab Tech. Lab*	*Tech. Excellence*
CSO4	*$700mn.*	*>1000*	*Division Lab (TD-Fac) 15%*	*15:70:15%*	*World Leader*	*BR Facility TD Facility*	*Tech. Excellence Ops. Excellence Customer Responsive*
CSO5	*$1.5Bn.*	*>1000*	*Central R&D Lab*	*1:10:89*	*Above Average*	*No*	*Tech. Excellence*
CSO6	*$30mn.*	*500-1000*	*Overall Co. R&D 60%*	*30:20:50%*	*Average (or Other)*	*Research Unit Acad. Unit*	*Tech. Excellence Customer Responsive- - ness*

subjects were equally divided among public and private sector organizations with one-third being academic institutions, labs or departments and the rest spread across medical engineering, information technology, telecommunications, electrical systems, and aerospace engineering. Collectively, they also used a number of different performance measurement and management frameworks.

Table 7-2 (above) presents certain other relevant organizational characteristics of the case study organizations (CSOs). Across the six case study organizations, the annual R&D expenditures ranged from $8 million to $1.5 billion, the sizes (in terms of number of employees) ranged from less than 20 to greater than 1000 technical and managerial employees. The case study subjects also represented an even mix of performance profiles with one of the six organizations describing itself as a world-leader, one of them describing itself as in top-quartile among its class of organizations, another three describing themselves as above-average and one as average (or other).

Four of the six organizations also had sub-organizational units (see Table 7-2) that allowed us to study the differential impact/level-of-implementation of strategy-performance-incentives systems on sub-organizational units doing different type of R&D work and having different missions to cater to. In terms of strategic posture as well, our case study subjects represented a well-balanced mix. Half of the organizations had two simultaneous strategic postures with two of the remaining organizations describing one and one describing three strategic themes as their strategic posture respectively.

TABLE 7-3: A PARTIAL LIST OF ORGANIZATIONS INTERVIEWED

Private/Corporate Labs
- Kodak Inc.
- Corning Inc.
- Intel Corp.
- France Telecom
- Dupont Corp.
- IBM Corp.
- Concept Labs
- W. R. Grace
- Air Products Inc.
- Omnova Solutions

University-Academic Labs
- Integrated Nano-Tech. Research Facility
- Brandeis University

Public Sector (non-Defense) Labs
- NASA
- National Energy Technology Laboratory
- Lawrence Berkeley National Laboratory
- Lawrence Livermore National Laboratory

Defense-related Labs
- Airforce Research Lab (AFRL)
- Aerospace Corporation
- The MOVES Institute

University-Academic Labs (contd.)
- Saint Louis University
- Cleveland State University
- Florida State University

In addition to the above case study organizations, we also conducted interviews with personnel from 32 R&D organizations of varying sizes and types to supplement the findings of the case study research. Table 7-3 (above) presents a partial list of organizations whose R&D leadership was interviewed for this purpose. This evidence has been used, anonymously, to supplement, corroborate, and validate the evidence gleaned from case study research.

In the next section, we begin by exploring the cumulative body of evidence gleaned through the six case studies. This analysis is organized along the key themes relevant to this investigation, namely, strategy-formulation, performance measurement, and incentives design functions. We also discuss the linkage between strategy-performance-incentives (SPI) systems, the achievement of "balance" between various dimensions of performance, and the prospects of implementing Balanced Scorecard-based SPI systems in R&D organizations.

7.3—THE CASE STUDY FINDINGS

In this section, we would draw upon the six individual case studies as well as 32 R&D organizations whose personnel were interviewed to present qualitative evidence relevant to the issues of relevance to us to this research, namely, strategy-formulation, performance measurement, and incentives systems in R&D organizations. In doing so, we would draw upon Yin's (1993) advice on theory building to aimed at testing a set of theoretical priors to achieve generalizability through theoretical or literal replication.

Our own theory draws upon and brings together several independent research streams in the areas that are relevant to this research. In instances where no such prior work or documentation exists (e.g. implementation of Balanced Scorecards in R&D organizations) we draw upon evidence from related literature in other settings and attempt to develop tentative hypotheses that may be subjected to scrutiny. These priors and hypotheses are (presented in Figure 3-4 are) reproduced in Figure 7-2 for a ready reference. We turn towards each of these and present the cumulative case study evidence.

This section is organized as follows. We start with addressing the perceptions of our case study subjects as well as practitioner interviewees vis-à-vis R&D performance measurement in general. We then examine how the strategy-making or strategic planning function is carried in these organizations, especially with an eye towards assessing whether it caters for multi-

FIGURE 7-2—THEORETICAL MODEL/PRIORS FOR CASE STUDY ANALYSIS (REPRODUCED FROM FIGURE 3-4)	
RESEARCH ISSUES	**HYPOTHESES & THEORETICAL PRIORS[89]**
Acceptability and Prevalence of Balanced Scorecard	***HS1.1:*** Balanced Scorecard (BSC) not prevalent among R&D organizations ***HS1.2:*** BSC Use: Private→Public→Academic ↓ (usage goes down); ***HS1.3:*** MAPMS prevalent among R&D organizations ***HS1.4:*** MAPMS Use: Public→Academic →Private ↓ (usage goes down);
Balance, Participation, Accessibility, and Transparency of Strategy-Making	***HS.2.1:*** Organizations using BSC/MAPMS take differential view of strategy ***HS2.2:*** Organizations using BSC/MAPMS with balanced/participative strategy-making systems report higher satisfaction and realize greater performance improvements. ***HS2.3:*** Organizations using BSC/MAPMS incorporate performance multi-dimensionality in strategy-making ***HS2.4:*** Organizations using transparent-participative strategy-making engage in strategic ("double-loop") learning
Balanced Scorecard and Multi-attribute Performance in R&D	***HS3.1:*** R&D organizations implement measurement systems that are "balanced" across dimensions of performance ***HS3.2:*** Organizations using BSC/MAPMS use cause-and-effect modeling to "measure the strategy". ***HS3.3:*** Organizations using BSC/MAPMS that incorporate qualities of HS3.1/HS3.2 report higher satisfaction and realize greater performance improvements.
Use of Incentives to Align Individual with organizational performance	***HS4.1:*** Organizations using BSC/MAPMS believe in ability of incentives to influence individual performance ***HS4.2:*** Prevalence of Incentives: Private→Academic → Public ↓ ***HS4.3:*** Perceived or Realized Effect of Incentives: # ("menu")of Incentives ↑, Intensity of Incentives ↑↓, Financial→ Professional ↑
Alignment in Strategy-Performance-Incentives Systems	***HS5.1:*** Organizations using BSC/MAPMS with better-aligned strategy-performance-incentives systems report higher satisfaction and realize greater performance improvements. ***HS5.2:*** Organizations using BSC/MAPMS that "cherry pick" components of the overall strategy-performance-incentives system for implementation report lower satisfaction and realize lower performance improvements.

[89] The notation for this column is as follows: "Private→Public ↑" means that as we move from private sector to public sector R&D organizations, the acceptability and prevalence of performance measurement decreases.

attribute character of R&D performance and is participative in nature. Next, we look at the conduct of performance measurement in these organizations. We use the generic performance measurement and management framework developed in chapter-3 as a lens to assess the procedural integrity of the measurement process as well as assess its impact on organizational performance itself. We also draw some insights vis-à-vis the implementation issues that might arise in the context of putting a measurement system in place. We examine the prevalence of incentives system and their perceived impact on individual performance. Finally, we discuss how the quality of linkage between strategy-making, performance measurement, and incentives systems affect the benefits realized from these systems.

7.3.1—Perceptions about R&D Measurement & Management Systems

The first of the series of issues that we were interested in was the perception about and acceptability of the idea of performance measurement and management within R&D organizations. The need for assessing the acceptability of performance measurement arises from a disagreement on this important issue within the R&D community. An often-expressed notion is that trying to measure R&D is a "fundamentally wrong-headed endeavor"[90] since R&D is a chancy process whose outcome cannot be predicted and hence measured. Another related argument deals with the distinction between research and development activities. According to this view, while it might be counterproductive to measure research, the case with development might not be so. In research, some argue, one would try to create the conditions to maximize the chances of a "surprise", while in development the opposite is true i.e. one would try to avoid surprises rather than welcome them. This makes the latter more amenable to managing and hence measuring than the former.

Despite these stereotypes, however, performance measurement and management is part and parcel of any management debate within the R&D community. It featured as the most pressing problem faced by R&D managers in the Biennial survey of the members of Industrial Research Institute (IRI) for two consecutive years (Ellis, 1993). We began our investigation with a set of priors about the perceptions of various types of R&D organizations towards performance measurement as gleaned through the relevant literature and our general understanding of the subject. We expected the acceptability and prevalence of R&D measurement to go down as we moved from private sector to public sector R&D organizations, from corporate to academic labs, and from development-oriented R&D activity to research-type activities. As we studied the six case study organizations, we met with several surprises that led us to fine-tune our theoretical priors.

We asked our respondents questions to assess the general organizational sentiment towards performance measurement. We also asked individual scientists, wherever possible, as to whether

[90] Personal conversation with Dr. Robert Spinrad, Ex-Director of Xerox PARC, June, 2003

they thought their performance was "measurable", whom did they think (i.e. themselves, their peers inside or outside the organization, or their managers) was best qualified to measure their performance, and whether they believed their organizations did a good job at doing so. We were also able to gain some insights into top-managements' perceptions about the utility of performance measurement and management systems by putting together pieces of evidence and anecdotes that related to the various types of performance measurement activities that were conducted, the importance that was attached to them, and how the resulting information was used within the organization.

On the whole, we found our priors on acceptability and prevalence of performance measurement and management systems to miss the true diversity of our target population. For example, we found our characterization of public/private, corporate/academic, and development/research dichotomies rather limiting for the purpose of this analysis. It was not sufficient to merely state that private sector organizations were more receptive to measurement of their R&D activities than public sector organizations or that corporate R&D was more receptive to measurement than academic R&D. One had to go a step or two deeper and explain why and how.

We found the reality to be far from the black-and-white distinctions across various sectors that we had hypothesized. There were significant gray areas within which a significant number—perhaps the majority—of the organizations seemed to operate. We did find some evidence to support our theoretical priors. For example, one of our case study subjects (CSO1)—a relatively small private sector R&D operation—had the most elaborate performance measurement system of all our case study organizations and another one, CSO6—an academic department of a public sector university—had the least elaborate measurement system. Yet, we also found counter examples to these sectoral distinctions.

Specifically, CSO3—a private sector R&D organization—employs a relatively more qualitative and individualized performance measurement philosophy than we would have anticipated on the basis of our theoretical priors. CSO3 manages to use a peer review-type system while operating in a cut-throat competitive environment with increasing pressure to justify its value to its corporate parent. The Director of CSO3 is confident that the performance measurement system, in its current shape and form, sufficiently suits the environment and mission of the organization. Similarly, CSO2—a public sector university—is pioneering the role of performance scorecards alongside the traditional academic peer-review type performance architecture—an innovation one would hardly expect to find in traditional academic settings. While none of the above counter-typical examples have matured enough or delivered yet to be considered unqualified successes, they nonetheless point towards the changing perceptions and trends that are cutting across the sectoral stereotypes.

If a generalization can be made of the cumulative evidence gleaned through the six cases, it is that one must factor in a host of other variables while trying to generalize across the diverse landscape of R&D organizations. While organizational type and type of R&D activity performed may be important—perhaps the most important—factors they are certainly not the only ones. Not only are there organizations and experimenting with new ways of measuring their performance, the general idea of performance measurement itself is gaining greater acceptability within R&D management circles. Increasingly, organizational types that were earlier thought to have a more reluctant posture towards measuring R&D are now doing so, albeit for reasons not anticipated by theorists and R&D practitioners alike.

For example, the traditional public sector lab that operated under no pressure to justify its existence to a corporate parent or demonstrate a return on investment for much of its documented history now operates under an environment that is increasingly demanding efficiency, effectiveness, and accountability of its operations. This is amply demonstrated in the debate surrounding the passage and implementation of GPRA and the drive to implement OMB-PART across the federal R&D portfolio. Public sector organizations are increasingly establishing performance measurement systems to comply with these requirements and are even going farther than what is being legislatively required of them. The example of CSO4—an R&D arm of a large public sector agency—that has experimented with the Balanced Scorecard methodology—for precisely the reasons one would expect private sector organizations to do so—is a case in point. Similar changes might be hypothesized to be at work in the academic sector, as is demonstrated CSO2's example.

In addition to the case studies, the experts/practitioners' interviews from the 32 R&D organizations generally confirm the sentiment evidenced from the cumulative case study evidence. While the performance measurement philosophies of each of these organizations varied with their unique mission requirements, expectations of the measurement process itself, and management styles, the act of performance measurement is becoming more acceptable and prevalent across the entire organizational landscape.

We did find some instances where individuals were not fully comfortable with the notion of performance measurement in R&D. For example, one of our interviewees believed that the R&D community is still "metric-averse" and another complained of "paperwork overload" (partly) due to the plethora of measurement schemes that had been used with little or no improvement in the desired outcome to show for them. The majority of our interviewees, however, believed that performance measurement, done wisely, was necessary, if not a welcome endeavor. Many of our interviewees pointed out that the R&D community was not metric-averse after all—provided they could relate to and see some value in the activity. The tradition of tenure system and peer review in academia was highlighted as the case in point. Others pointed towards the increasing use of Six Sigma and project-based measurement approaches in industry and TQM in the government as a

sign of changing trends. We also found considerable support for the proposition that sectoral (i.e. public/private, corporate/academic, research/development) differences are one of the several factors that determine an organization's perception towards measurement of R&D performance.

7.3.2—Strategy-Making in R&D Organizations

Within the realm of strategy-making, we were interested in the manner in which the process is carried out within R&D organizations, and more specifically, whether the strategy-making process itself was participative, transparent, and multi-dimensional in its intent and implementation. The literature on the Balanced Scorecard describes, in some depth, the virtues of organizational strategy being the foundation on which a performance measurement system needs to be built. *"Measuring the strategy"*, in this view, is the key objective of a well-intentioned performance measurement architecture (Kaplan and Norton, 1996, 2001). This literature also credits the awareness and ownership of organizational strategy, achieved primarily through the ability to participate in the strategy-making process, as a key ingredient of the much-publicized performance breakthroughs achieved by organizations that have implemented a Balanced Scorecard.

TABLE 7-4– STRATEGY MAKING AMONG CASE STUDY ORGANIZATIONS

CSO#	STRATEGY-MAKING SYSTEM	DESCRIPTION OF STRATEGY-MAKING SYSTEM	OTHER KEY ATTRIBUTES
CSO1	Management-by-Objectives-type	Top-down, bottom-up	Multi-attribute, partially consultative, multiple sources of guidance
CSO2	Compact Planning	Participative/consultative	Top-down, bottom-up, multi-attribute, decentralized implementation,
CSO3	Indus. Peer Review	One-to-one consultations	Research agenda-dominated, researcher interests/freedom, iterative,
CSO4	Bureaucratic	Top-down, reactionary	Multiple sources of guidance, glacial changes in strategy, partially consultative
CSO5	Traditional Strategic Planning Process	Project-driven	Flexible to meet project needs, partially consultative,
CSO6	Academic Peer Review Process	Collegial/consultative	Permanence/stability,

The six case study organizations that we looked at represented a diverse set of strategic planning approaches ranging from traditional top-down strategic planning to one-on-one consultations between stakeholders to the use of Balanced Scorecard-type participative and consultative processes. The salient features of these processes are presented in Table 7-4 (above). While three of the six case study organizations had Scorecard-type in place, only one (CSO2) came close to integrating its strategy-making process with its Balanced Scorecard. CSO2's strategy-planning process was an adaptation of a participative strategy-making approach to the more traditional and collegial consultation process in academia.

In this top-down and bottom-up process, referred to as *Compact Planning*, CSO2's top leadership identified institution-wide priorities while the subordinate administrative departments and academic units (and their constituent sub-units) came up with initiatives, milestones, and performance targets to contribute towards these institution-wide objectives. Compact planning encouraged the administrative and academic departments to identify their own department-level goals and objectives in addition to those prescribed for the overall institution. While, on closer examination, the reality on ground of how the compact planning process was actually implemented differed substantially from the concept on paper, we observed considerable learning within CSO2's various departments (or units) that is a sign of progress in the right direction.

Another dominant approach—found in at least two of the six case study organizations—was to organize the strategy-making process around projects or a research agenda. In fact we also found this to be a fairly common strategic-planning model among the organizations interviewed for the purpose of this research. Within the two case study organizations, we found two different shades of such a process being in place. Operating style and size had translated into substantial differences in the strategy-making processes at these organizations.

CSO3—a small Bell-Labs-style private sector R&D facility—that greatly valued the independence and innovative ability of its scientists and engineers and used a relatively hands-off approach to management, adopted a one-to-one approach to strategic planning. This list of R&D projects or the research agenda was developed through a consultative process that took place between individual researchers (or principal investigators) and the unit directors in a consultative and collegial manner. The second of the project-driven strategic planning processes was found at work in CSO5—a large private sector organization engaged in aerospace R&D. CSO5 had different types of strategic planning processes dealing with everything from small exploratory projects to large multi-billion government contracts. These ranged from localized consultations between researchers and research managers (or directors) to large organization-wide decision processes that required the approval of the top-management to go forward. Depending on the size of the project and the maturity of the

science or technology, researchers' (or scientists') roles varied from active decision-makers to indirect and semi-passive implementers.

7.3.2-1—The prevalence of participative and multi-attribute strategy-formulation in R&D— ***N***otwithstanding the different types of strategic planning approaches observed, the two unmistakable trends that emerge from our six case studies are the growing popularity of both consultative-participative and multi-attribute approaches to strategic planning. There is a growing realization of the importance of creating an organization-wide ownership of the strategy within the R&D management community. While the practice, in its current form, falls far short of what would be an R&D-equivalent of a Balanced Scorecard Best Practice, it is certainly moving in that direction. On querying whether R&D managers believed that involving everyone in the strategy-making process could improve the efficiency and effectiveness of the process itself, we found considerable openness but also some trepidation and caution.

In a number of our case study organizations (i.e. CSO4, CSO5, CSO6) R&D leaders were gradually opening up the strategy-making function to the bench-scientist over and above his or her traditional role and involvement as custodians of organizational capability and knowledge. Several of our respondents also highlighted the fact that not all researchers may be interested in participating in the strategy-making process and hence one must always keep room for the eccentric genius who likes to be preoccupied with the intellectual challenge of his work and does not care much about organizational strategy. In one of the organizations (CSO4) we found evidence that a section of the employees were completely unaware of the existence of a strategy-making or performance measurement initiative. Other organizations—CSO2 being one of them—are inducing their employees to pay attention to organizational strategy by substantial material rewards and raising the stakes for non-involvement. For example, CSO2 and CSO5 link budgetary allocation and new initiatives funding with their strategic planning processes thus inviting participation from even those skeptical of getting involved in such exercises.

There also seems to be an almost universal consensus on a multi-attribute view of performance. Three of the six organizations in our case study sample are implementing their versions of a Balanced Scorecard that requires a multi-attribute approach not only performance measurement but also strategy formulation. Here, once again, there is a considerable gap between the ideal practice (per Kaplan and Norton 1996, 2001) and the reality on the ground. With the exception of CSO2 none of the other organizations use a multi-attribute approach to their strategic planning processes ex-ante, although they all claim to do so, ex-post. For example, CSO4—being a government agency—has a fairly elaborate and bureaucratic strategy-making process that draws upon multiple sources of guidance with the desire to achieve balance across multiple dimensions of performance as one of several concerns. CSO5's strategic planning system, on the other hand, is heavily influenced, ex-ante, by a project-driven approach to strategic planning. For both these organizations using the

strategic planning process to drive a balance across multiple performance dimensions is an after-thought rather than a driving consideration.

In addition to and apart from the three organizations that implemented a Balanced Scorecard, a multi-attribute thought process was also a common theme within the strategy-making processes of other case study organizations. CSO1 used an elaborate Management-by-Objectives strategic planning process that identified performance targets for each of the five key dimensions of performance emphasized by the corporate top-management. CSO3 used an elegant bi-dimensional approach to strategic planning and performance measurement and CSO6 used the standard academic emphasis on research, teaching, and service as key dimensions of performance. In some ways CSO1 and CSO3, that do explicitly claim to implement a Balanced Scorecard, seemed to do better than CSO4 and CSO5 (described earlier) in incorporating, ex-ante, a multi-dimensional view of performance in their strategy-making processes.

Clearly, the evidence gleaned from the six case studies falls short of validating the theoretical priors on the acceptance of participative, multi-attribute strategy-making approaches within various sub-sectors of R&D organizations. Private sector entities do not seem more or less likely to accept and implement participative strategy-making than public sector ones, as evidenced from the examples of CSO1, CSO3 and CSO5, on the one hand, and CSO2 and CSO4, on the other. While academic organizations (e.g. CSO2, CSO6) have traditionally been more collegial and participative in their approach towards strategic-planning, and corporate labs (e.g. CSO1) more hierarchical and less participative.

Finer distinctions like missions, operating style/philosophy, history and other factors do matter in determining an organization's openness to adopting a participative strategy-making process. We see the effect of some of these factors at play within CSO5 that shows multiple flavors of strategy-making approaches within the same organization. The lesson to be drawn here, as before, is that broad dichotomies like public/private, corporate/academic and research/development may not be very helpful, in and of themselves, as predictors of an organization's tendencies towards a particular strategic planning posture. One must look a level deeper to reach a finer level of understanding.

7.3.2.2—Participative strategy-making, individual motivation, and organizational performance— Finally, we looked for evidence of an impact of participative strategy-making processes on individual and organizational performance. Prior to undertaking the case studies, we had hypothesized several possible mechanisms through which a participative strategy-making process might affect individual or organizational performance. These included, among others, better information about strategy, a sense of ownership of the strategy, a greater sense of fairness-transparency-empowerment, a greater focus on strategy, and a better strategy itself. We looked for evidence to support or refute each of these hypothesized mechanisms. While we found no

conclusive concrete quantitative evidence either way, what we did gather from several of our interviewees were statements expressing satisfaction or dissatisfaction with the state of affairs as far as their role in the strategic planning process was concerned.

For example, At CSO1—an organization that implements a fairly traditional top-down strategic planning process—we found some dissatisfaction with the manner in which organizational priorities were determined and employees were asked to achieve stretch targets year-after-year. There was a sense among our interviewees that employees were over-worked and their intellectual curiosities under-nourished as a result of a top-down corporate-like approach to research priority setting. In another organization (CSO3) that adopted a more consultative approach to research priority setting and strategic planning, we encountered a more relaxed and satisfied sentiment. Here, scientists and engineers were made amply aware of what was expected of them and prospective employees were only hired if they were comfortable with being judged against these performance expectations and targets.

The above anecdotes, however, only present suggestive evidence. We cannot claim to make the link between satisfaction with the status-quo and improvements in individual performance without resorting to elements of theory that were not covered in this analysis. All three of our case studies that had implemented a Balanced Scorecard had done so relatively recently with the result that the system was not mature enough to begin showing major signs of performance improvement. However, with the exception of CSO2 that showed some positive signs, the Balanced Scorecard itself—at least in the manner in which it had been implemented across three of our six case studies—does not seem to make a difference in the quality of the strategic planning process or the ownership of strategy itself. CSO2 comes closest to adopting a participative approach to developing a Balanced Scorecard and while the initial signs (vis-à-vis the adoption by and learning of employees) are encouraging, whether or not this would translate into performance improvements of the kind claimed by the founders of the Balanced Scorecards remains to be seen. These issues of temporal maturity and quality of implementation are classic problems in measuring the impact of management frameworks within organizations (Hackman and Wageman, 1995).

Qualitatively, the debate within the majority of the case study organizations and the 32 R&D organizations where personnel were interviewed hovered not around whether or not individuals—bench scientists and technicians—should be involved in the strategy-making process in the first place but rather on what was the right approach to get the momentum going and what proportion of the R&D employees actually paid any attention to organizational strategy. The perceptions of R&D managers and practitioners that we interviewed varied considerably on those counts. The actual practice of involvement of scientific personnel in strategy-making process ranged from highly informal (e.g. requiring every employee to answer a simple question about his/her satisfaction with the direction and management of the facility to town-hall meetings, newsletter-based discussions,

roundtable discussions, all-hands meetings and celebration of an innovation week to ideas solicitations) to highly formal (e.g. agency-wide strategic planning exercise, management retreats, and appointment of independent commissions) highlighting this variation in perceptions.

Most of our interviewees believed that there was considerable anxiety among the scientific and technological workforce vis-à-vis their role and contribution to the organizational goals. "Majority of the employees want to see the 'big picture'. Only a minority remain unengaged", was the response of the provost of a large private university. Even for the marginally unengaged, it was believed, that one only had to show the employees how everything that they did connects with the big picture to inspire them to pay attention. Others were more skeptical of the possibility of involving everyone in the strategy-making process and the good that might come out of it. "We do not want to close off opportunities for involvement but not everyone can get involved in strategy-making all the time", was the response of the Research Director of a large chemicals company.

Here, and at several other organizations (e.g. a large telecom lab and a large IT lab) strategy-making was considered the domain of the organization's top-management, albeit with some advisory involvement of the subject-area experts. "Scientists have their own biases [that must be balanced against organizational interests represented in this process by the top-management]", was a comment we heard from the Research Director of the telecom lab. Whether or not scientists have biases, the organizational interests and goals need to be incorporated in the process. An open question then for each of our interviewees was whether this is done using a top-down or a more participative approach with all participants having some influence on the overall outcome. Each of these organizations settled somewhere in the middle of the continuum of that best suited its own circumstances and style of management.

The key challenge here, in addition to creating opportunities for employees to get involved in the strategic planning process, is not only to communicate adequately the expectations from such an exercise (ex-ante) but also the results of the process (ex-post). We found anecdotal and quantitative evidence of the fact that several of our case study organizations failed to do well on that benchmark. On several instances (e.g. in at least one of our case study organizations) employees showed lack of awareness of the organizational strategy or the process put in place to solicit their input in the same. One of the case study organizations conducted formal employee surveys to assess perceptions and found a weak employee understanding of strategic issues. On the contrary, however, organizations that did well to communicate expectations and strategy (e.g. CSO3) or made efforts to consolidate and continually improve on this account (e.g. CSO2) seemed to do well in inspiring a positive attitude among their employees towards organizational goals and objectives. Many of these attitudes, however, had not resulted in quantifiable improvements in individual and organizational

performance primarily because the initiatives had been in place long enough for the effects to take place.

The sum-total of the cumulative evidence on the two issues of concern to us, therefore, is that while there is an unmistakable shift towards a more participative and multi-dimensional strategy-making process, the degree and manner in which these are implemented vary considerably. There appears to be no agreed answer on the optimal level of employee participation is concerned. Also, contrary to the prescription of the Balanced Scorecard literature which encourages the use of performance multi-dimensionality as a lens through which organizational leaders and employees must look at strategy, multi-dimensionality comes merely as one of the several considerations, more often than not, in our case studies, as an after-thought, in the strategic management process. While the trend seems to be in the right direction, this could explain why none of the organizations that we studied could achieve significant and noticeable differences in performance that are attributable to empowerment of employees through their participation in strategy-making.

7.3.3—Performance measurement & management systems in R&D organizations

To what extend do the performance measurement and management systems at work among R&D organizations recognize multi-dimensionality of performance? Using the Balanced Scorecard as a reference model, we were interested in whether: a) R&D organizations used a Balanced Scorecard (or similar frameworks), b) among those R&D organizations that did use the Balanced Scorecard, whether its use and perceived benefits resembled those in more familiar corporate settings, and c) among those organizations that did not use a Balanced Scorecard, whether the alternate measurement framework delivered the kind of performance improvement normally associated with the Balanced Scorecard literature.

The first thing that we learnt was that R&D organizations did use multi-attribute approaches in general and Balanced Scorecard-type approaches, in particular. The realization that performance was a multi-dimensional phenomenon was universal among our case study subjects whether or not the organization used a performance measurement philosophy that explicitly catered to it. Here again, we found our theoretical priors seriously lacking, as we failed to discover any sectoral differences between organizations that did or did not take a multi-attribute view of performance measurement. We were fortunate to study three organizations that implemented a Balanced Scorecard among our six case study subjects—one each public, private and academic. Each of the remaining three organizations also took a multi-attribute view of performance by identifying explicit and specific performance categories deemed critical for its performance. In addition to the Balanced Scorecard, our case study organizations also used a variety of other performance measurement frameworks ranging from industrial and academic peer reviews to management-by-objectives to six sigma and the stage-gate process. Table 7-5 (below) presents the salient features of the performance measurement frameworks at work in each of our case studies.

TABLE 7-5 – PERFORMANCE MEASUREMENT (PM) SYSTEMS AMONG CASE STUDY ORGANIZATIONS				
CSO#	**PM FRAMEWORKS USED**	**STRATEGY & PMS LINKAGE**	**PERCEIVED IMPACT ON PERFORMANCE**	**OTHER COMMENTS/FACTORS**
CSO1	MBO, Stage Gate, Six Sigma Methods,	Partial/Unclear	Slightly Positive	Financial and non-financial metrics, performance categories do not follow through,
CSO2	BSC/Dashboards, Peer Review	Partial/Unclear	No Impact	External communication tool, 8-performance categories,
CSO3	Peer Review, Individual Appraisals	Instances of mis-alignment/Clear	Very Positive	Clarity of expectations, simplicity (2 performance categories, 2 metrics encapsulate everything)
CSO4	Balanced Scorecard Individual Appraisals	Partial/Unclear	Slightly Positive	3-performance categories
CSO5	Balanced Scorecard, MBO	Partial/Unclear	Very Positive	3-performance categories,
CSO6	MBO, Academic Peer Review	Instances of mis-alignment/Clear	No Impact	Overlapping/counter-productive 7-performance categories

7.3.3.1—Measurement "overload" and the use of multiple performance measurement frameworks—The majority of our subjects used at least two (and up to four) performance measurement frameworks simultaneously. Many times these different frameworks operated in distinct organizational domains (e.g. separate frameworks for measuring individual and organization performance). Nevertheless, there were instances, and understandably so, where these frameworks interacted and exchanged data with each other. We hypothesized that the organizations that would do a good job at either minimizing the number of simultaneous performance measurement frameworks so as to minimize the "measurement overload" or creating a seamless and purposeful interface between two (or more) such frameworks would tend to do better than those that do not. We found several examples of the above phenomenon among our case study organizations.

CSO1 presents a particularly interesting example of what we believe are telling signs of a measurement overload. At CSO1, while management-by-objectives served as the central

measurement philosophy, several other measurement frameworks also contributed to the overall organizational performance measurement system. Many tens of performance metrics were being collected within the organization—the majority used rather cursorily if used at all. An individual seeking guidance on his or her own performance could easily drown in several overlapping measurement systems and a plethora of metrics without getting to the right set of metrics that he or she could realistically control and respond to. The inter-linkages between some of these measurement frameworks were unclear, at best. As we probed our interviewees about the whether or not they saw any marginal value in having additional measurement frameworks in place. In one instance, for example, a six-sigma-type process review was incorporated on top of an already existing and highly stringent MBO process. The reason advanced for doing so was that the corporate parent of CSO1 had such a system and asked all its subsidiary operations to do as well.

A counter example is that of CSO3 which, by virtue of its clearly defined mission and well-articulated objectives and performance requirements, has a simple performance measurement framework that boils down to monitoring two performance metrics for all employees, namely, number of publications (quality-adjusted) and contribution to technology transfer. Well-defined targets for each of these metrics at each level of seniority provide clear expectations of performance to each individual within the organization. Managing and measuring performance within CSO3, therefore, becomes a matter of monitoring these metrics and providing timely feedback to all concerned. The simplicity and elegance of the performance measurement framework makes it an apt mechanism for self-evaluation, peer pressure, and feedback within the organization.

7.3.3.2—The quality of linkage between performance measurement & strategic planning systems— Another issue of relevance to the present discussion that was made conspicuous by the cumulative case study evidence is that of the linkage between the strategic planning and the performance measurement systems. Simply speaking, in order to "measure the strategy" effectively, the performance measurement architecture must directly follow from the strategy itself. A failure to do so would either entail measuring the wrong things or wasting precious management time measuring and strategizing about things that really did not matter to the performance of the organization. Among the six case study organizations we found considerable variation in the manner in which performance measurement systems linked back to the strategic planning systems or failed to do so. These variations ranged from systems (e.g. at CSO1, CSO2, and CSO5) that were only partially aligned with the strategic planning systems to those (e.g. at CSO3 and CSO4) that were more or less well-aligned with the latter. We, however, did not find a textbook implementation of a Balanced Scorecard where the performance scorecard follows through directly from the strategy itself.

Leaving aside for a moment CSO3 and CSO4—where the discrepancy was of a minor nature, we focus attention on the more serious cases of misalignment. CSO1 and CSO2 provide examples of one

type of misalignment where the strategic planning process itself comprises an amalgam of several different approaches—often working in tandem with each other—producing a complex, at times, even confusing picture of what an organization is supposed to do. In CSO1, for example, there were many such sources of strategic advice, namely, vision, mission, values, goals, and processes which, although intended to supplement each other, did not necessarily relate to or follow from each other. In CSO2 as well, we had several different sources of strategic guidance, namely, the compact planning documents and the academic peer review process. While both these organizations had strategic planning processes that used the same language as the performance measurement framework and were conducted in a manner that reflected concerns for the same set of constructs, the presence of too many sources of strategic planning necessitated that the linkage between the two could be only partial, at best.

In CSO4 and CSO5, one the other hand, the misalignment between the performance measurement architecture and the strategic planning process was of a qualitatively different character. Both these organizations used Balanced Scorecards as their primary performance architectures. Each also identified three dimensions of performance on their Balanced Scorecards. Here, as against in CSO1 and CSO2, not only did the various strategy formulation processes did not talk the same language but also as a result of the former there was a clear disconnect between the strategy-formulation processes and the performance measurement systems. Thus, the three performance dimensions that are central to the structure of these organizations' Balanced Scorecards appear to almost feature as an after-thought in their strategic planning processes.

7.3.3.3—Lessons from the implementation of the Balanced Scorecard in R&D organizations— Notwithstanding the above issues, our case studies did illustrate some preliminary examples of the implementation of Balanced Scorecards in R&D settings. They provide a first hand view of the various issues and problems that might arise during such an exercise and what might be different from the Balanced Scorecard in corporate settings. The three of the six case study organizations that implemented a Balanced Scorecard did so in slightly different ways. CSO2 started the use of a Balanced Scorecard (it calls it a "performance dashboard") primarily as a device to convey its performance to its external stakeholders. Their performance dashboard has evolved, quite predictably, over the years and improved in terms of its coverage of critical dimensions of performance. It is only after at least a couple of years of its use as an external tool that CSO2's management has begun to understand and emphasize its importance as a tool to monitor organizational performance internally. CSO4 and CSO5, on the other hand, have their Balanced Scorecards primarily focused on internal performance, as per the Balanced Scorecard literature.

All three organizations made credible attempts to cascade their scorecards further down the organizational hierarchy—an effort that was not always equally well-received within various parts of their organizations. In the case of CSO2, for example, the scorecard was readily adopted by the

administrative units but not so well by the academic departments. One can contrast this with CSO4 where both the research and the product development divisions willingly signed on to the cascading process and developed their own specific scorecards. These organizations provide clear evidence of the importance of top-management support, communication, and gradualism on the acceptance of the scorecard deployment process within the lower echelons of the organizational hierarchy. In all three organizations, however, the scorecards were first conceived and later refined (e.g. identification and refinement of performance perspectives and individual metrics) in a manner that resembles its development in the corporate settings.

Another major challenge faced by R&D organizations implementing a Balanced Scorecard—as amply illustrated in the examples of CSO2, CSO4 and CSO5—is their inability to link their scorecards with strategy satisfactorily. We have already discussed this issue in some detail above but here we would like to point out a more fundamental problem encountered in at least two of the three case study organizations, namely, the lack of a credible focused strategy itself that could reflect the changing needs and realities of the competitive landscape.

Specifically, all three attempts to implement the Balanced Scorecard were focused on measuring the performance and overall health of the organization in a more or less stable fashion rather than geared towards "measuring the strategy"—a difference of emphasis perhaps that is important to the essence of the Balanced Scorecard. One can hypothesize why this might be so. One of the reasons, we believe, is the relatively long cycle times in an R&D setting as compared to the corporate setting. Take the example of a change in R&D strategy of an aerospace company where a decision is made to invest in and develop the capacity to produce a new kind of stealth technology. Even though this might be a significant strategic change, it may take several years—maybe even a decade or more—before it could firmly sink-into various parts of the company's R&D and non-R&D operations. Contrast this with the decision of an oil-marketing company to provide a facelift and a different selling proposition to its several hundred convenience stores—a change that would take a few weeks to execute and could begin influencing the bottom-line performance within months. The extent to which long cycle times and feedback loops are a characteristic of R&D settings, all R&D organizations are likely to encounter such issues with a need to adjust the expectations of how a Balanced Scorecard is expected to deliver.

7.3.3.4—The impact of implementing a Balanced Scorecard on organizational performance— Whether or not a multi-attribute performance measurement system makes a difference to the organizational performance itself is a difficult question to answer, under the circumstances, since all six case study organizations seem to take a multi-attribute view of performance. Given the above problem with identification, we may narrow down the query and try to address whether or not the use of the Balanced Scorecard seemed to make a differentiable impact on organizational

performance itself. If we look at the self-described perceived impact, we find no identifiable difference between the organizations that implement a Balanced Scorecard and those that do not. Of the three organizations on each side of the divide, one each perceives the measurement approach to have a "very positive", "slightly positive", or "no impact" on the performance itself. This we believe is not a non-finding. The absence of a discernable difference in performance pre- and post-Balanced Scorecard—or the "performance breakthrough", so to speak—goes counter to our expectations of how a Balanced Scorecard is supposed to work. We are fortunate enough, however, to look a level or two deeper in the organizations to ascertain the causes for this finding.

One probable explanation may be the lack of maturity of the Balanced Scorecard efforts. A first-order cursory analysis of the ground reality seems to support such a hypothesis. As the Balanced Scorecard efforts mature, the organizations seem to report greater satisfaction with them. CSO2—with a relatively nascent effort—reports the least satisfaction while CSO4 and CSO5—with 6-8 year-old efforts report much greater satisfaction with them. This, however, may only be a part of the reality. Even the more mature of the efforts (i.e. CSO4 and CSO5) do not seem to achieve the kind of "performance breakthroughs" that are often credited to successful implementations of the Balanced Scorecard in the literature.

To a major extent, however, the reasons for these shortcomings may lie in incomplete or faulty implementation of the framework rather than the Balanced Scorecard per se. For example, of the three organizations that implemented the Balanced Scorecard none employed cause-and-effect performance models and strategy-maps that are instrumental in simplifying and conveying organizational strategy. We address this issue in some detail in section 7.3.5 (below.) Suffice it so say here that in none of the three cases that we examined did the Balanced Scorecard seem to inspire the kind of strategy and performance focus that it is often credited with. In all fairness, while there were instances where the use of scorecards set into motion a cooperative-synergistic dynamic within the organization (e.g. between various units of CSO2) or where the scorecard-based performance information was used to identify performance bottlenecks and improve upon them (e.g. in CSO4), these effects were largely localized and fell well short of producing the "laser-like" focus on organizational performance and strategy-execution. A deep organizational commitment to a full-scale implementation has a much greater chance of delivering than a half-hearted and skeptical approach to implementing the Balanced Scorecard.

The cumulative evidence from the 32 organizations interviewed for the purpose of this research does not counter the above conclusions. These organizations varied considerably in terms of the types of performance measurement approaches they used and how they used them. For example, we found that while majority of the organizations did take a multi-dimensional view of performance, the performance dimensions may not necessarily correspond one-to-one with the ones prescribed by the Balanced Scorecard literature. Industrial R&D labs, for instance, routinely

measured R&D through a mix of input, process, output, and outcome measures—a multi-dimensional view of sorts—and academia routinely measured its performance through a mix of measures relating to research, teaching, and public service but not all of these organizations paid equal attention to employee morale, customer satisfaction, or innovation management. Many of the interviewees justified this by one of the two arguments, namely, a Balanced Scorecard category being subsumed by one that they did measure (e.g. "the fact that we are achieving our output/enrollment targets means that our customer/students is/are satisfied") or the Balanced Scorecard category being measured through informal means ("we do not need to measure employee morale, we can guess that through their emails"). While there might be some truth to this argument, it clearly falls short of the Balanced Scorecard ideal and hence needs a closer examination.

The 32 organizations interviewed for the purpose of this research, like our six case studies, also used a variety of performance measurement approaches. The most common among these were Six Sigma methodologies, Total Quality Management, Management-by-Objectives, Stage-gate and other project-based approaches, and peer review. We did not find many other examples (with the exception of a couple) of the use of the Balanced Scorecard among our interviewees. This clearly hints at the fact that while Balanced Scorecard is one of the approaches used by R&D organizations it is not as popular as some of the others mentioned above. [91]

The use of specific measurement approaches seemed to cluster around sectoral lines with six sigma and stage-gate/project-based approaches being more popular among industrial R&D labs, TQM/MBO among public-sector labs, and benchmarking and tenure system/peer review among academic labs. The reason for these trends had more to do with the ability of a particular technique to match with the needs of a particular type of R&D activity than with the generalized willingness (or lack of it) towards performance measurement within these sectors per se. A stage-gate or project-based measurement approach aggregated at various levels of the organization fit well with the predominantly development-oriented activity within hierarchical organizations of the industrial labs. Six sigma-type techniques owed their popularity to the use of concepts of scientific methods (e.g. design of experiments)[92]. TQM's popularity within public sector labs lent itself to a government-wide thrust to adopt this methodology but it also, along with MBO and peer-review, better suited the type of work and mission considerations of public sector R&D labs. Same is the case with the use of benchmarking, tenure system, and peer review in academia.

---

[91] Our case study design required an over-sampling of organizations implementing a Balanced Scorecard.

[92] One of our respondents, a mid-sized chemicals company, had this to say about the popularity Design for Six Sigma (DFSS) techniques at their campus: "Scientists love tools and Six Sigma provides them with some. They use experimental design techniques to build new molecules. They do not teach that in school. We have been successful in learn some important skills through the use of six sigma approaches here". The Director of another Chemicals company concurred with that when he added: "Six Sigma is big here. R&D people must understand how to eliminate unnecessary variation and design-in quality in their experiments"

The cumulative evidence from case studies and organizational interviews on the substance of the questions posed at the beginning of this section seems to suggest that while multi-dimensional performance measurement systems are common in R&D, the dimensions may only have an imperfect correspondence with the performance dimensions of the traditional Balanced Scorecard. Also, barring some exceptions such as the use of input, process, output, and outcomes metrics, the performance multi-dimensionality in R&D took a form (e.g. multi-stakeholder or key performance indicator model) that was different from the cause-and-effect performance model of the Balanced Scorecard literature.

An even stronger finding pertains to those organizations that claimed to have implemented a Balanced Scorecard. We found considerable shortcomings in the extent and quality of their implementation efforts. These organizations implemented only certain features (e.g. performance multi-dimensionality, cascading) of the Balanced Scorecard without regard to the completeness of the overall framework. Conspicuously missing were the cause-and-effect performance models, the strategy-maps, double-loop learning, and a strong link with organizational strategy. The effects on performance itself, therefore, seemed to have been seriously compromised due to faulty and incomplete implementation. This inconclusiveness of the evidence demands a more detailed investigation in future studies of the Balanced Scorecard.

7.3.4—Incentives Systems within R&D Organizations

"Do incentives matter" is an oft-asked question within the R&D management community. The community is deeply divided between those that believe that scientists and engineers are motivated by intrinsic factors and those that believe that extrinsic factors do matter as well. The reality perhaps lies somewhere between the two extremes. With regards to incentives structures within R&D organizations, we were specifically interested in three features of our subjects' incentives systems, namely, the mix of incentives (financial vs. professional; high-powered vs. low-powered), their linkage with performance measurement, and their impact on individual performance and organizational performance itself. Due to the vast nature of the subject of incentives and rewards, we limited our focus to very pragmatic consideration of incentive systems as an aid and supplement to the strategic-planning and performance measurement activities described above. What follows is thus only a partial—and pragmatic rather than philosophical—treatment of the subject.

7.3.4.1—The prevalence and types of incentives systems in R&D—Unlike the strategic planning and performance systems (discussed above), the sectoral distinctions were much more pronounced in the incentives and rewards systems of organizations (Please see Table 7-6). They were also much more consistent with our theoretical priors (see Figure 7-2). As far as the prevalence and intensity of incentives is concerned, we found a more or less clean and consistent picture. The incentives

structures at two of the three private sector organizations (i.e. CSO1 and CSO5) were more high-intensity than either the academic or public sector labs. These organizations generally emphasized financial incentives and the sharing of fruits of performance with their employees. The presence of stock options, bonus plans, and executive compensation arrangements gave strong and credible signals of the managements' intent to share the benefits of performance organization-wide. One exception to this rule was CSO3 that showed a mixed incentive environment with features of both academic and corporate settings. On the one hand, for example, it depended on intrinsic sources of motivation (e.g. novelty, freedom, and intellectual challenge) to inspire performance and, on the other hand, it offered to handsomely reward superior performance (e.g. through a reward for "extraordinary" effort to commercialize technology to be paid to the recipient in the "currency of his choosing").

TABLE 7-6– INCENTIVES SYSTEMS AMONG CASE STUDY ORGANIZATIONS

CSO#	SECTOR / TYPE OF R&D	COMPENSATION SYSTEM	FINANCIAL INCENTIVES	PROFESSIONAL INCENTIVES	LINKED TO PERFORMANCE/SCORECARD
CSO1	Private / Development	Strong/Performance-based	Cash Rewards, Bonuses	Time/Money to attend Conferences	Performance targets for all, No scorecards
CSO2	Academic / Research	Weak/Cost-of-living (COL) & equity-based	Cash rewards, Royalties etc.	Autonomy, Time/$ for Conferences	Performance targets for some, Not linked to scorecards
CSO3	Private / Research	Medium/Performance-based	Cash rewards, Royalties etc.	Autonomy, Recognit--ion, Promotion,	Performance targets for all employees, No scorecard
CSO4	Public / Devlopment	Weak/mostly COL and YIS-based	Cash rewards, Royalties	Recognition, visibili--ty, $ for Conferences	Performance targets for some, Not linked to scorecards
CSO5	Private / Development	Strong/Performance-based	Spin-off, Cash, Stocks, Group	Recognition, visibili-ty, promotion	Performance targets for all, Not linked to scorecards
CSO6	Academic / Research	Weak/Performance- and COL-based	Cash, Royalties etc.	Recognition, visibili-ty, $ for conferences	Performance targets for some, No scorecard

The one (and only) public sector organization (CSO4) in our sample of case studies presented a relatively weaker incentives environment. The absence of a financial bottom-line and limitations of civil service regulations were cited as possible reasons. The predominant philosophy here was to compensate for a lack of strong financial incentives by focusing on professional rewards and recognition and by providing flexibility and choice instead. This organization provided a "menu of incentives" from which employees could pick and choose what best suited their interests and

circumstances. A similar approach was adopted by another public sector university (CSO2), again due primarily to civil service limitations and financial constraints. Both these types of organizations tended to heavily depend upon intrinsic factors of motivation to inspire their employees to perform better. This practice was found to be prevalent much more so in academia (CSO2, CSO6) than in the public sector organization (CSO4) and much more so together—the academia and the public sector combined—than their private sector counterparts (CSO1, CSO5).

In addition to the positive incentives, we also encountered the prevalence and strong effects of negative incentives (or lack of them). Two of the six case study organizations raised the issue of negative incentives in somewhat different ways. CSO3 claims to benefit from the possibility of a negative incentive (i.e. laying-off of an employee in the case of a failure to perform) in its attempt to create a performance culture. Employees not only clearly know what is expected of them but also what is likely to happen if they do not come up to these expectations. The research director at CSO3 identified specific instances where talented well-meaning individuals voluntarily quit the organization after failing to meet these expectations.

A similar dynamic is seen at work among universities where the tenure process serves the same purpose. The tenure system, however, is a double-edged sword. Not only does it encourage younger faculty to perform better before they have attained tenure, it also perhaps makes the more experienced faculty more complacent once they have attained tenure. We also saw, in CSO4, a glimpse of what might happen in the absence of a negative incentive in public sector—a public sector equivalent of post-tenure incentives environment. The case in point is the widely perceived erosion of performance in public sector bureaucracies among mid-level bureaucrats where civil service regulations make it increasingly hard to terminate ones employment.

The cumulative case study evidence also illustrates an important factor in determining the incentives structures within public, academic, and private-sector settings. The organization's ability and willingness to trade-off individual autonomy and freedom with high-powered financial incentives to a large extent, seems to determine its philosophy towards the incentives system. Several exogenous and endogenous factors, namely, mission, organization, operating philosophy, and civil-service limitations, contribute towards it. Academic labs trade-off financial incentives for autonomy while private sector do the reverse providing more pay to their employees but taking away, to varying degrees, their autonomy for choosing what to work on. These cases also illustrate the notion of self-selection on the part of employees—scientists and engineers—to seek employment in an organization that offers an incentive bundle that balances their taste for intrinsic vs. extrinsic factors.

7.3.4.2—The impact of incentives and rewards on individual performance—Did incentives—financial or professional—seem to make a difference in terms of improvements in actual

performance of employees? The general consensus on the subject, as gleaned from the six case study organizations, slightly tilts towards a rather modest expectation on this front. While all the organizations—across the board—talked about and emphasized the limitations of financial incentives in influencing performance—a feeling that was strongest in academia. "You cannot make a researcher work smarter by paying him or her more", was a typical reply. The solution: hire the better and more talented of the people, give them freedom and space of choosing their own research agendas, acknowledge their contributions, and reward them modestly without creating inequity within the organization. Professional incentives and recognition seemed to be in considerable favor with most of the respondents.

In spite of the general feeling, however, organizations—in both public and private sectors—tried to provide incentives to their employees that ranged from performance bonuses and stock options to special awards, plaques, and medals. Some of these, especially private sector organizations, seemed to believe that incentives did make a difference. One of our interviewees at CSO1—a high-intensity incentive environment—argued that performance bonuses did work by raising the stakes of bad performance and providing the necessary feedback to the employees. Many others, however, believed that while financial incentives may not influence performance ex-ante in any significant way ("nobody starts to work smarter just to win a pot of money") it might be a good way to acknowledge good performance ex-post. Several of our respondents cited the example of the Nobel Prize—the ultimate reward for accomplishment in science and technology—and wondered how many scientists and researchers' work was influenced by a realistic assessment of their chances of winning one. In most instances, they seemed to argue, Nobel Prize winners were motivated to do interesting and useful work and the recognition (and money) came only as an ex-post acknowledgement of their achievements.

For others still, providing incentives—financial, but also professional—was a way to retain employees in an increasingly competitive marketplace of researchers. They believed that researchers were smart and very aware people who knew what their skills were worth in a fair market valuation and did not want to be taken advantage of in a financial sense. "Since everybody seems to do it, we too must follow the pack", epitomizes the crux of this sentiment. On the whole, the evidence vis-à-vis the actual impact of incentives on performance was rather mixed. We heard well-articulated arguments and some anecdotal, though not conclusive, evidence on each side of the debate.

The 32 organizations we interviewed allowed us to seek and entertain a wider range of opinions and experiences vis-à-vis the role of incentives in improving R&D performance. They, however, did not seem to change the qualitative nature of the conclusions drawn above. If anything, these organizational interviews seemed to accentuate the deep divisions on this important issue, not only between the various sectors but also within the same sector. While traditional approaches and

incentive structures dominate the scene, universities are increasingly experimenting with ways to incentivize performance beyond the traditional approaches discussed earlier.

"Perhaps, times have changed", was the response of a University administrator that we interviewed. "When I joined academia several decades ago, I took a pay cut. Today, young faculty members are more aware and show a deep interest in incubators, royalties, and start-up and consulting opportunities". Another one of the universities that we looked at was contemplating putting in place a bonus system that would allow a researcher to earn up to 50% of his or her salary as a performance bonus, provided he or she can bring in enough external research dollars to more than compensate for his or her salary. Notwithstanding these efforts, however, a tight resource environment, concerns for equity across departments, civil service regulations, and a traditional mind-set that emphasized intrinsic factors of motivation presented substantial hindrances to this move.

Within the corporate sector as well, the additional evidence gained from the 32 interviewees seemed to have a dampening influence on what we had learnt from our case study organizations. While the private R&D organizations still dominated the other two sectors in intensity of the financial incentives, we also heard counter arguments and examples that emphasized the need for balance. "An incentive system that only focuses on financial rewards is likely to lead to dysfunctional behaviors", asserted one of our interviewees. Generally, organizations experimented with a whole range of financial and professional incentives and provided their employees with the flexibility to pick and choose from them thus validating our earlier conjecture vis-à-vis the popularity of the "menu of incentives" approach, not only among public sector organizations that used it most but across all types of R&D organizations.

As with the case study organizations, tenure process and civil service regulations—specifically, the dampened incentive to perform and the inability to fire an individual failing to perform as per expectations—were seen as major impediments to productivity and performance. One of our interviewees—a provost of a large public sector university—suggested early-retirement schemes and life-long performance compacts as possible remedies to the situation. Another—a director of a public sector lab—uses employment contracts that fall outside the realm of civil-service regulations as his key recruitment vehicle.

7.3.5—Correspondence between the Generic Performance Measurement and Management Framework and the Strategy-Performance-Incentives Systems of Case Study Organizations

Having looked at the three components of the strategy-performance-incentives systems in isolation with each other, we now look at them as a broader strategic management and measurement system along the lines similar to a Balanced Scorecard of Kaplan and Norton (1996, 2001). The generic

performance measurement and management framework developed in chapter-3 would serve as a reference for this discussion. This framework outlines a set of generic principles for putting together effective performance measurement and management systems.

7.3.5.1—Performance multi-dimensionality in theory and practice—The first and foremost issue raised by this generic performance measurement and management framework is that of performance multi-dimensionality itself. Specifically, we were interested in determining whether or not R&D organizations take a "balanced" and multi-dimensional view of organizational performance and whether they do so in an implicit or explicit manner. At the very outset of our case study, we made a conscious decision to look at organizations with both types of performance measurement systems, namely, those that appeared to take a multi-attribute/dimensional view of performance and those that did not. Through this conscious approach, we hoped to learn about how these different performance measurement philosophies played out in the R&D setting.

The cumulative case study evidence on the matter appears to allow the following generalization: all R&D organizations that we studied, irrespective of whether or not they explicitly adopted performance measurement approaches that were built around a multi-attribute philosophy of performance measurement, seemed to take a multi-dimensional view of performance. The multi-attribute character of the strategy-performance systems of some of our case study organizations is graphically illustrated in Figure-7-2. Clearly, all of our case study subjects either explicitly (CSO2, CSO4, and CSO5) or implicitly (CSO1, CSO3, and CSO6) incorporated a multi-dimensional view in their strategy-performance systems with the former group doing so through the use of a Balanced Scorecard-type philosophy while the latter group incorporating it in more traditional MBO/peer-review type measurement approaches.

We have also tried to assess the similarity of these multi-attribute approaches with the 4-part Balanced Scorecard framework by superimposing the latter on top of each one of the strategy-performance architectures. While neither of the strategy-performance architectures have a one-to-one correspondence with the performance perspectives of the Kaplan and Norton Balanced Scorecard, one can clearly identify significant overlaps. All performance architectures seem to

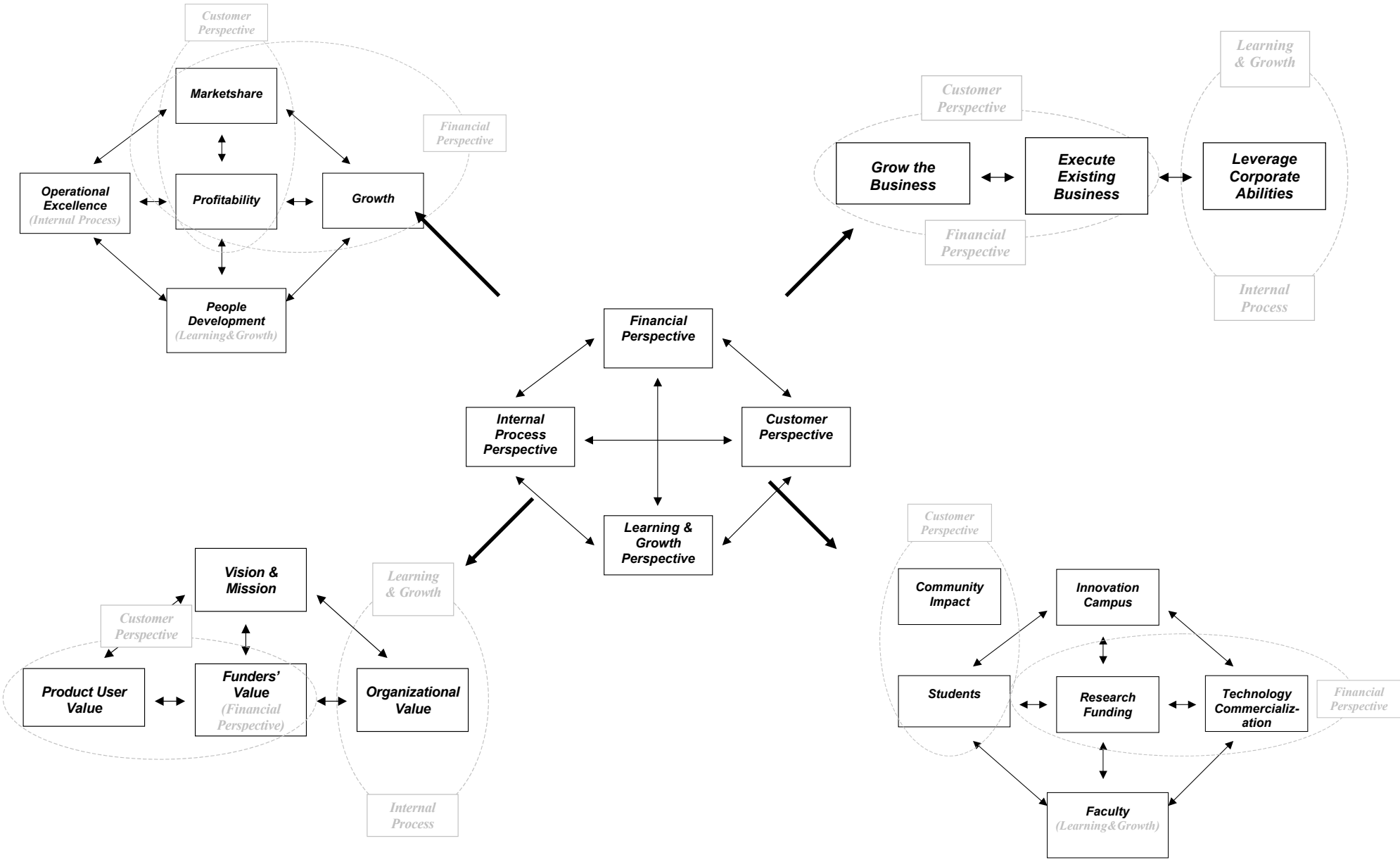


Figure 7-3: Performance Multi-Dimensionality in R&D—Cross Case Comparison

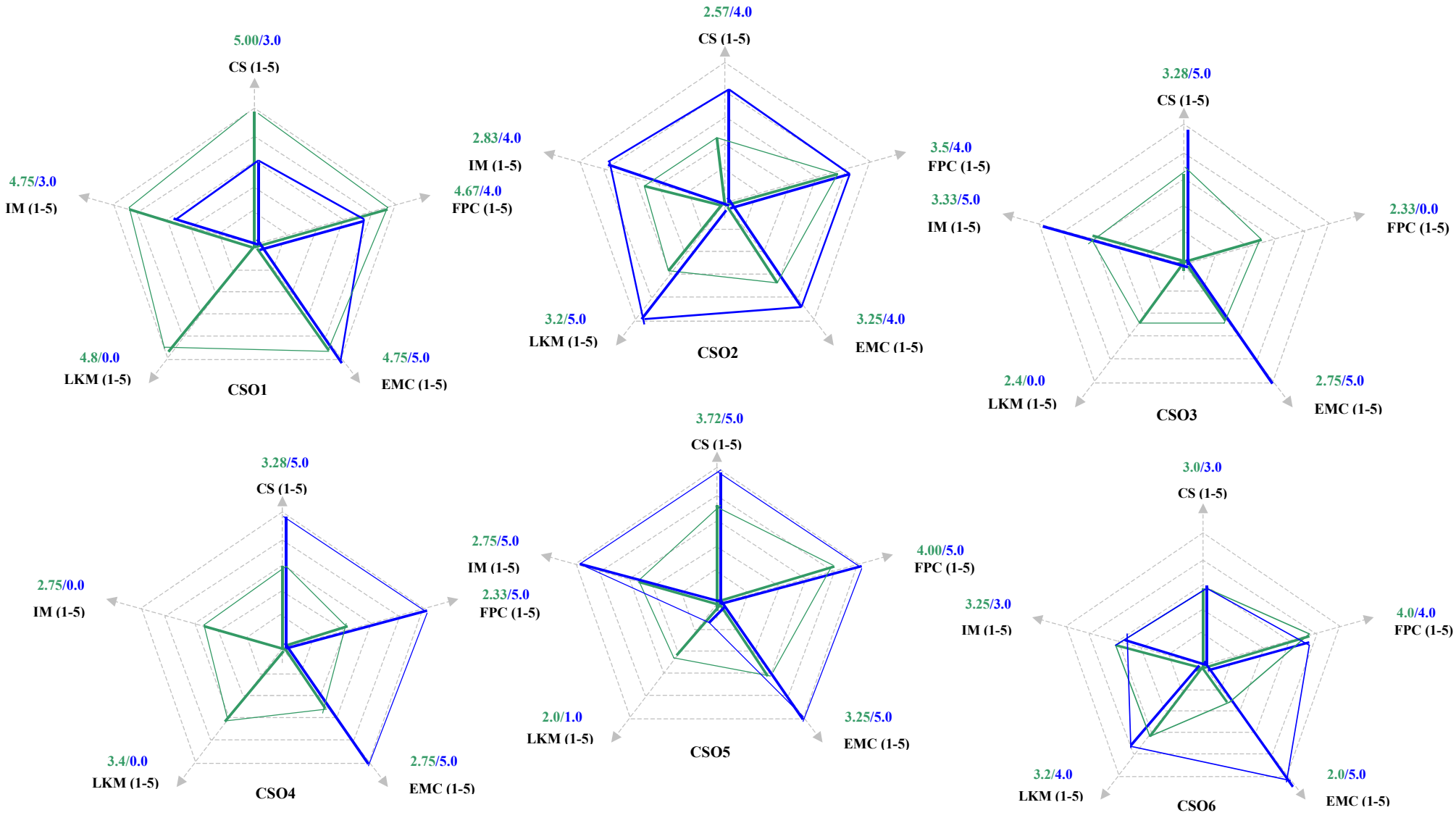


Figure 7-4: Cross Case Analysis – The Degree of Balance In SPI Systems

incorporate certain performance dimensions of the Balanced Scorecard, albeit in a rather imbalanced manner. For example, an organization's strategy-performance architecture (top-right, figure 7-3) may have several dimensions dealing with the financial perspective while at the same time only a single dimension catering for learning and growth and internal business process perspectives—a discrepancy that defeats the very essence of the Balanced Scorecard.

We also used another reference model—our own generic R&D Balanced Scorecard—that was developed and described in chapter-3 as a means for comparing "balance" across various dimensions of performance. We did this by indirectly asking each of our six case study organizations if they used some (or all) of the five key dimensions of performance contained in that model. We also provided them a detailed questionnaire of sub-categories that comprised the five broader categories of performance and asked them to identify whether and to what extent were each of these emphasized within their organizations. The results are summarized graphically as spider charts in Figure 7-4.

The scores (1-5) depicted by the blue lines (or blue spider webs) describe whether the organization in question explicitly deemed a particular performance dimension (e.g. innovation management, learning and knowledge management etc.) as critical to its performance. The green lines describe a cumulative measure of whether or not the organization emphasized upon certain sub-dimensions of each of the five performance dimensions—irrespective of whether it explicitly did so.

The disparity between the scores represented by blue and green spider webs is quite evident. It reflects a significant finding that can be generalized across all of the six case study organizations. Even though half (three of the six) case study organizations did not explicitly emphasize all five dimensions of performance deemed critical in the reference model, they in fact did so in an implicit manner. Another interesting observation for the astute observer might be that five of the six organizations (excluding CSO1) thought they paid more emphasis to a particular performance dimension (per blue spider webs) than they actually did (per green spider webs) while only one (CSO1) thought it paid less emphasis than it actually did—a case, perhaps, of "over measurement". This finding points out towards an "optimism bias" among organizations vis-à-vis the comprehensiveness or coverage of their strategy and performance architectures.

7.3.5.2—Other key differences between theory (generic performance measurement Model) & practice—Coming back to the discrepancies between theory and practice, we encounter a general lack of emphasis on critical structural features of the Balanced Scorecard in actual implementations. Most noticeably missing is the use of cause-and-effect models that form the central core of the Balanced Scorecard philosophy. This is virtually a universal finding across all six of our case study organizations—quite striking as it is, but even more so for those that claim to implement a Balanced Scorecard (i.e. CSO2, CSO4, and CSO5). While each of the six case study organizations, with minor exceptions, did seem to incorporate metrics belonging to each of the four performance perspectives,

they did not make any attempt to link these in a cause-and-effect model that would allow strategic learning of the sort hypothesized by Kaplan and Norton (2001). While we did find some evidence of organizational learning, these were isolated instances that—in the absence of a cause-and-effect model of organizational performance— would not allow "what-if-ing" or hypothesis testing of any significant nature. This, we believe, is a major aspect that needs the attention of organizational leaders and managers alike.

Another important element on which the case study organizations varied considerably was their ability to communicate strategy to their employees. In the traditional Balanced Scorecard literature this is accomplished through the strategy-map that simplifies the organization's strategy for its employees and presents it in an easy-to-understand graphical format. The effectiveness of the strategy-map, however, rests on the use of cause-and-effect logic. We found some suggestive evidence of the fact that our case study organizations did not do very well at communicating strategy to their employees.

TABLE 7-5– GENERIC PERFORMANCE MEASUREMENT & MANAGEMENT FRAMEWORK & CASE STUDY ORGANIZATIONS (CSOS)

CSO#	IMPLEMENT BALANCED SCORECARD	MULTI-DIMENSIONAL	CAUSE-AND-EFFECT LOGIC	STRATEGY MAPPING	CASCADING	STRATEGIC MGMT SYSTEM
CSO1	No	Yes, Indirect	Partial	Partial	Yes, Strong	Partial
CSO2	Yes	Yes, Indirect	No	No	Yes, Weak	Yes, Growing
CSO3	No	Yes, Direct	Partial	No	Yes, Strong	Yes, Strong
CSO4	Yes	Yes, Indirect	No	No	Yes, Strong	Partial
CSO5	Yes	Yes, Indirect	No	No	Yes, Strong	Weak
CSO6	No	Yes, Indirect	No	No	No	Partial

Organization size and mission simplicity and coherence played a critical role in determining the size of this dysfunction. Smaller organizations (e.g. CSO3) or those with a simple and coherent mission (e.g. CSO6) tended to do better than larger organizations with diversified missions (e.g. CSO4 and CSO5). Telling evidence of this was found at one of the case study organizations (CSO5) where the majority of employees scored worst on questions relating to the understanding of strategy, structure and purpose of the Balanced Scorecard on an annual employee survey conducted to ascertain attitudes, perceptions, and knowledge about various organizational constructs.

The Balanced Scorecard--by virtue of the fact that it did not incorporate a cause-and-effect logic—did not seem to make a substantial difference to organizations' ability to communicate strategy. Our

case study organizations generally did well to cascade the top-level metrics down the organizational hierarchy. However, even when organizations were able to cascade the strategy and performance systems, the overall "big picture" of organizational strategy still remained beyond the grasp of an individual—thus only affording the possibility of performance improvements that were localized rather than global and stove-piped rather than synergistic.

Finally, we were interested in whether or not and how our case study organizations use their strategy-performance-incentives systems as the cornerstone of their strategic management systems. The motivation for this undertaking arises from what is generally perceived as the gradual maturation of the Balanced Scorecard from a performance measurement system to a strategic management tool (see Kaplan and Norton, 1992, 1993, 1996, and 2001 for details). There is a parallel understanding that major gains in its acceptability within the organization and some of the key benefits of the Balanced Scorecard arise from its use as a centerpiece of the strategic management system rather than a performance measurement system. Only when the Balanced Scorecard is used as a basis for strategic conversations within the organization and accorded that importance by the top-management, do employees further down the organizational hierarchy begin to realize its importance and get involved in implementing the strategy as espoused by the founders of the Balanced Scorecard.

Consequently we were interested in whether or not Balanced Scorecard and other types of performance measurement approaches integrate with the broader strategic planning systems. We find some evidence support such a maturation process, although not every organization does so with equal degree of success or promise. Specifically, the example of CSO2—a public sector university that is experimenting with a Balanced Scorecard-type performance dashboard—stands out as one moving in the right direction. CSO2's compact planning process and the performance dashboards have, over the years since their inception, become the centerpiece of its budgetary allocation process. The organization's top management has also signaled the importance of the compact planning process by tying a certain portion of discretionary and new-initiative funding with the initiatives identified on departments' compact plans. The result has been lesser resistance to the idea of compact planning itself and a greater focus on strategy and resource utilization.

Still other organizations do not seem to be as successful as CSO2 in linking their Balanced Scorecards or performance measurement systems (in general) to their strategic planning processes. The result, in this scenario, is often a strategic planning process that is disjoint and isolated from, or only loosely coupled with, the actual day-to-day decision-making and resource allocation. Needless to say that wherever we did find a greater coordination between the performance measurement and the strategic planning systems, we sensed a greater sense of empowerment and initiative among those using it.

Whether or not this greater sense of ownership translated into gains in actual performance remains an open question. While we found some anecdotal evidence that points towards the fact that R&D organizations whose strategy-performance-incentives systems show greater resemblance with the generic framework tend to do better in terms of performance, it is largely inconclusive for reasons best attributed to a lack of identification. More simply, we do not have, in our case study sample, even a single organization whose strategy-performance-incentives system shows a one-to-one correspondence with the general principles identified in the generic model and hence no reference against which to compare the performance of the R&D organizations against.

While we do find some evidence of the fact that organizations that incorporate greater number of structural features tend to report greater satisfaction with their strategy-performance-incentives architectures, it would be a leap of faith to present that as an evidence of demonstrable performance improvements. We, therefore, fail to convincingly validate or reject our hypothesis as defined in Figure 7.2.

7.3.6—Implementation Challenges and Bottlenecks

Before we close this section, we briefly refer to some of the broad lessons and findings in the realm of implementation challenges and bottlenecks. While this discussion neither is nor is intended to be a conclusive one on the subject, it does highlight some issues that we encountered during our case study exercise. That the process of implementation is as important as structure of the Balanced Scorecard itself is a well-known finding across much of the case study literature on the Balanced Scorecard.

The most important and often-cited implementation challenge vis-à-vis a management or measurement technique in general and the Balanced Scorecard in particular is the initial skepticism among employees that must be overcome through clarity in communication and intent and consensus building within the organization. This challenge assumes a more acute dimension in the context of an R&D organization. Scientists and engineers are, by nature and training, prone to not accept anything without proper facts and verification. They are also very skeptical of "management's" attempts to interfere in their work and try to organize their workplaces and work patterns. These feelings are further aggravated by the fact that many have dealt with several management frameworks in the past and have found them to be of little use or actual relevance to their work.

Operating in the backdrop of this mistrust and skepticism, it is important to introduce a new performance measurement methodology—like the Balanced Scorecard—with great caution and care and appropriate emphasis on communication and consensus building. In the words of one of our interviewees, the proponents of a new measurement/management paradigm need to answer the classic question from the standpoint of a bench scientist/engineer: "What's in it for me?".

Organizational leaders and managers and idea champions must understand their the bench scientists/engineers may not share their objectives and motivations and try to address their concerns as they push forward with the implementation of the measurement/management framework.

Many of our interviewees highlighted that, contrary to the common belief, scientists and engineers are not necessarily averse to performance measurement. In fact, many argued, that barring certain initial resistance that is expected at the introduction of anything new, they welcome the measurement of their performance as long as it has a credible and meaningful relationship with the results of their work. They are skeptical of and averse to the idea of "measuring for the sake of measuring only" without any real and meaningful affect on research output or performance. It is under these circumstances that they view performance measurement merely as an added burden on them rather than an exercise of any utility.

We found at least a couple of dominant ways in which our case study organizations tended to counter this initial skepticism and resistance. One of the approaches used (e.g. by CSO2) was to tie significant stakes with participation in the strategy-formulation and performance measurement activities. The gradual acceptance of CSO2's compact planning process, even by those who initially criticized it as a mere management fad, is one such example. CSO2 overcame this initial criticism from and reluctance of the researchers/scientists by linking it with the budgetary allocation process—thus signaling not only the seriousness of its own commitment but also raising the stakes of non-participation. Another important factor found in CSO2, but also elsewhere, was the ability to demonstrate some initial benefits to the skeptical bystanders.

A somewhat related point is the emphasis on organizational communication and the need to clarify not only the new "rules of the game" but also its goals and expectations. CSO2 encountered such problems early on in the process as several of its sub-units were not able to understand either the intent or working of the compact planning process. Failing to make it work to their advantage, several of these sub-units became disillusioned with its utility and ran the risk of joining the category of its critics within the organization. Unhampered by this but also empathetic to the situation, CSO2's management took effective steps to further communicate the intention and working of the process and, having trained the requisite staff, repeated the process again. Quite expectedly it managed to attract greater participation and less skepticism from its various sub-units. CSO3 also lays great emphasis on communication as a means to clarifying performance expectations across its organization. These organizations tended to do better in terms of getting their people involved in implementing the performance measurement systems.

Another related matter is the simultaneous use of several strategic planning and performance measurement frameworks within the same organization without regards to the marginal value of adding another framework on top of the others. We encountered this in several of our case study organizations, the most prominent being CSO1 where a six sigma process was recently introduced

and a 360-degree feedback[93] process was being planned on top of several well-established processes, namely, a stage-gate process, an MBO process, and an individual performance appraisal model called standards of leaderships (SOL) that were already in place. Under circumstances such as these, not only is one likely to find greater resistance to the idea of implementing yet another management framework but it is also difficult for the organizations' management to convincingly demonstrate the marginal benefit of the newly added frameworks. It is important for R&D leaders and managers to carefully think about these issues before embarking upon such an initiative.

Finally, another important implementation lesson that stands out from the cumulative case study evidence is that of selective application of certain aspects of the management and measurement framework without regards to its effect on the quality of the overall framework. While it might make sense to do so as a part of a gradual approach to implementation (i.e. pick the low hanging fruit first and score some easy gains to convince the employees before coming to the more painful parts), it is not advisable in and of itself. The various aspects of the generic performance measurement and management framework, as described in chapter-3, fall into place only when used together. For example, in the absence of a multi-attribute organizing framework, there can be no cause-and-effect model of performance, and without a cause-and-effect model, it is impossible to create a meaningful strategy-map. Failing to implement the various aspects of the Balanced Scorecard may result in seriously limiting the possibilities of achieving a performance breakthrough. We saw this dynamic at work in majority of our case studies.

7.4—Discussion and Tentative Conclusions

Several interesting and useful themes emerge from the above analysis, complementing, supplementing, and at times clarifying some of the results obtained in the concurrently carried out mail survey of R&D organizations. From a methodological standpoint, if there is one message made clear by the case study component of this research, it undoubtedly is that *"details do matter"*. Every organization's story is unique and one cannot understand and appreciate its strategy-performance-incentives systems without looking at it in great deal of depth and detail while at the same time putting it in its contextual surroundings. For example, two organizations claiming to implement a Balanced Scorecard maybe doing quite, if not fundamentally, different things by adapting the framework and selectively apply features that best suits their own needs and expectations. They might also be achieving different levels and types of benefit from these exercises thus making it challenging for observers to compare their respective gains to each other or to a specific reference standard. In many ways, these organizations, by virtue of their inherent differences and those of the expectations and motivations of their leaders, are implementing different things. We lose this diversity in contexts, motivations, and approaches when we try to force the standardized character

[93] The 360-degree feedback is an appraisal system premised on the idea that an individuals performance must be assessed by those higher and lower to him or her in the hierarchy as well as the peers.

of a mail survey on these subjects. The case study brings out and makes conspicuous the importance of the contextual details in a way none of the other methodologies perhaps can.

From a substantive standpoint as well, a number of important findings that emerge from the foregoing analysis have a strong consequence to the core subject of this research. First, based on the case study and interview evidence, we can conclude that the notion of performance multi-dimensionality is inherently natural to R&D organizations—perhaps even more so than their commercial counterparts. This is a fairly robust finding across multiple case studies (literal replication) and in line with strong theoretical underpinnings (theoretical replication). All six of our case study subjects and numerous others that were interviewed for the purpose of this study adopted a multi-dimensional approach to thinking about organizational performance. These approaches ranged from key performance indicators (e.g. CSO4, CSO6) to stakeholder scorecards (e.g. CSO5) to multi-dimensional academic and industrial peer review processes (e.g. CSO2, CSO3) to the incorporation of multiple dimensions in the more traditional strategic management paraphernalia e.g. goals, vision, and organizational core values (e.g. CSO1). We find a strong theoretical basis for why this might be so. R&D activity and organizations, unlike their commercial counterparts, have not had the luxury of an easily definable bottom-line thus forcing them to depend upon multiple dimensions and measures of performance. This finding can be construed as supporting our hypothesis HS1.3 (Figure 4-4).

Second, R&D organizations take a diffused rather than differential (a la Porter, 1996) view of organizational strategy. None of our six case study organizations took a differential view of strategy, especially during times where no organizational threats or crises existed. The strategic conversations at the majority of our case study and interview subjects were devoid of differentiation (in the marketplace of research) and pitched more toward promoting organizational longevity by doing a little better or more of the same. One exception is CSO2—a public-sector university-- that carefully identified its competition in academic and research services and developed a strategic plan to move to a desired state.

One possible explanation for the above finding may be a lack of sophistication, among R&D leaders and managers, in strategic planning concepts and processes. It is likely that the notions of differentiation in the market place of research services and differential strategy are so foreign to R&D leaders and managers that they fail to comprehend it in its intricate details. Reger et al. (1998) talks about the notion of a "range of convenience" for individuals' cognitive processes where they are more comfortable using a particular construct and beyond which they could find it hard to comprehend or apply the same construct. Seen in this context, "competition" and "differentiation" can be two such constructs that researchers and academics are quite at home with in the individual context yet so hard to comprehend and apply when thinking in organizational terms.

It is also likely that several organizations felt, some justifiably so, immune to external and internal competition and hence perceived little need for engaging in strategic planning of the kind we are talking about. This, too, is partly borne out from the case study evidence whereby several organizations had very long strategic planning cycles. It is quite likely that loose and fuzzy performance feedback loops coupled with long R&D cycle times provide R&D organizations a certain degree of freedom from the disciplining forces of market, and possibly, in some cases a false sense of freedom, at least temporarily, from disciplining forces of any kind thus providing them with the luxury of not needing a differential strategy. Contributing to and confirming this view was the finding that majority of our case study and interview subjects seemed to measure, some more explicitly than others, the notion of organizational health rather than the execution of a focused strategy.

An alternate explanation for the above finding may also reside in the inherent character of the R&D organization itself. The process of research and discovery is a chancy process that requires an organization to hedge its bets. Pelz and Andrews (1966), for example, have identified diversification as an important characteristic of better performing individual scientists and researchers. Diamond (1998) has documented the career consequences of failed research projects for individual scientists. Pelz and Andrews' high-performers hedge against these possibilities by pursuing several related lines of research within a single promising research area in the hope that at least one or two of these projects would lead to significant findings. This is a "diffused" strategy at the level of an individual and it is likely that R&D organizations take a similar view of organizational strategy. Which possible explanation or what combination explain the lack of differential strategies among R&D organizations is potentially a fruitful area for research. Regardless of which of these explanations reflect the truth, however, this finding results in our failure to confirm an important hypothesis, HS2.1 (Figure 7-2).

Third, while all of our case study subjects use a multi-dimensional framework for measuring organizational performance, this framework generally does not run uniformly throughout the organizational fabric and its strategic management system. For example, in several case studies (e.g. CSO5, CSO6) we noted that the use of performance metrics along multiple dimensions come as an after thought rather than a by-product of the strategic planning process. In these organizations performance multi-dimensionality was merely one of the several simultaneous sources of strategic guidance. The multiplicity of frameworks used to guide strategic planning could potentially result in confusing the messages received by an organization's employees regarding the organization's strategic direction and their own contribution to it. There are exceptions. Those organizations that have a coherent and uniform strategic management framework running through all aspects of major organizational functions seem to do better than others in focusing employees' energies in support of organizational missions.

Without a doubt, we found a general tendency towards more participative and transparent strategy-making but considerable variation in the extent and quality of such employee participation. On the whole, the weight of evidence is on the side of better satisfaction, if not improved results, among organizations that paid a greater emphasis to participative and transparent strategy-making processes. This partly confirms hypothesis HS2.2 (Figure 4-4). We fail to convincingly identify better outcomes in terms of performance improvements—although there were some hints of that in CSO2—that could be attributed to employee participation in strategy-formulation. We also found some evidence of double-loop learning (i.e. "what-iffing" and questioning of assumptions about organization's performance) among organizations that implemented more participative strategic planning structures and processes, however, this was often an exception, rather than a norm. This, we believe, indicates the inclination of individuals and organizations rather than a general outcome that R&D organizations seek as a result of implementing more participative strategy-formulation purposes. It points towards considerable unexploited potential of participative strategy-making functions that are increasingly becoming a norm in R&D settings.

Fourth, the case study evidence portrays an organizational landscape that remains deeply divided on the issue of incentives. Virtually all organizations, while acknowledging the importance of intrinsic motivational factors (e.g. challenge and nobility of work, freedom and autonomy), adopt a series of extrinsic incentives to motivate performance and align individual interests with organizational goals. We found partial confirmation of our hypothesis (HS4.2) regarding sectoral stereotypes of the prevalence of incentives in R&D organizations. Private sector R&D labs seemed to have the most extensive set of financial and professional incentives. Public sector entities seemed to trade-off incentive intensity with flexibility (i.e. number or "menu" of incentives) while academic labs seemed to trade-off incentives with individual autonomy and flexibility of choosing ones work.

With the exception of private sector labs, our case study subjects highlighted caution in the use of high-intensity incentives and even questioned the notion that researchers' performance can be influenced. "You cannot make people work smarter by providing them more money [in the pocket]", was a typical response. The general consensus was in the favor of the ex-post, rather than ex-ante, value of incentives as a positive ("feel good") factor and more importantly as a confirmation of the value and worth of ones work. In this sense, even financial incentives served to substantiate and reinforce the intrinsic and semi-extrinsic (professional) incentives for individuals. The value of incentives as resources to perform research (e.g. start-up grants) or further one's professional pursuits (e.g. conference attendance) was, however, universally acknowledged. There was also some evidence of an externally rather than internally driven incentives regime with competition for talent and concerns for cross-organizational equity driving organizations' policies towards financial pay-outs to their scientific and research staff.

On the whole, while we did find confirming evidence in support of our hypothesis that organizations believe in the ability of incentives to influence individual performance (HS4.1), the

exact mechanism of how this is achieved is open to differing interpretations and theories. Generally, it is believed that an indirect approach that attempts to align individual motivations with organizational goals provides superior results to a more direct broad-brush approach of trying to make individuals work smarter.

Fifth, while three of our six case study organizations claimed to implement a Balanced Scorecard, each chose to modify it to suit its own needs, contexts, and motivations. In arriving at a modified framework, these organizations selectively deployed some elements of the strategic management system and missed several others. To be fair, this constitutes a fairly generic and well-documented organizational problem and is not unique to the Balanced Scorecard per se. Hackman and Wageman (1995) describe in some detail the analytical and practical issues arising from selective application of some elements of an organizational management framework. Applying Hackman and Wageman's (1995) tri-part evaluation framework to the above problem, we are inclined to conclude that all three of our case study subjects that claim to have implemented the Balanced Scorecard have done so in the manner seriously deficient, from both structural and process standpoint. The omission of a number of structural features (described above) that constitute the conceptual core of the Balanced Scorecard leads to a lack of finding of several process outcomes (desired features) thus seriously undermining both the credibility of these exercises.

Quite predictably, therefore, none of the organizations that we studied reported the kind of extraordinary results documented in the relevant literature. While we found some evidence to suggest that organizations with well-designed and better-implemented strategy-making and performance measurement architectures seem to report higher satisfaction with both the process and outcomes of such efforts, these results fall considerably short of expectations from a successful implementation of a Balanced Scorecard.

CHAPTER # 8

CONCLUSION & RECOMMENDATIONS

In the final chapter of this dissertation, we consolidate the disparate pieces of evidence from our multi-pronged methodology and put together a consolidated picture. We translate these findings into a set of recommendations for an organization seeking to implement a Balanced Scorecard-type strategy-making and performance measurement architecture. Finally, we discuss some philosophical and methodological implications of our findings and identify future directions for research.

8.1—BALANCED SCORECARDS IN R&D

We premised this research on an assumption that Balanced Scorecard or similar scorecards-based multi-attribute performance measurement approaches were not as widely used in R&D as they had been used in the corporate and increasingly coming to be used in the non-profit and public sectors. The primary hypothesis for the lack of adoption of the Balanced Scorecard in R&D was that the process of value creation, objectives and expectations, or the type of individuals that engaged in research and development pursuits makes it less conducive to the sort of measurement techniques and philosophy that the Balanced Scorecard represents.

Our review of the literature on R&D measurement revealed a lack of consensus around determinants of what constituted quality and "good" performance in R&D. While we did find a growing consensus within the theoretical and practitioner communities on adopting a systems approach to thinking about R&D performance we found the literature to be lacking in the sophistication available in other similar streams of literature, e.g. healthcare and governance. This was also apparent in the manner in which several authors have recently made a call for adopting the Balanced Scorecard in R&D without adequately undertaking the requisite development and adaptation of the approach for R&D settings. The task that we faced, therefore, was to look at the Balanced Scorecard, on the one hand, and R&D performance measurement, on the other hand, and attempt to reconcile the two literatures. This required looking at the problem from two lenses. A critical evaluation of the Balanced Scorecard, as applied across various settings—especially R&D settings—asked what potential benefits and advancements over the state-of-the-art, is the Balanced Scorecard likely to bring if applied to an R&D organization. Alternatively, by bringing a fairly extensive literature on R&D management and measurement to bear on the problem, we assessed whether R&D was indeed different and if so whether those differences would entail that a Balanced Scorecard-type approach may not be an effective way to measure and influence performance. In essence, we asked the question: What aspects, assumptions, and elements of the Balanced Scorecard can we apply that are likely to deliver results within R&D settings and how might we modify the framework to suit the unique requirements of R&D organizations? We briefly summarize these findings below—following roughly the order of our research hypotheses (identified in chapter-4).

8.1.1—Acceptability and prevalence of the Balanced Scorecard in R&D—Contrary to our earlier belief and hypothesis on the subject (HS1.1), Balanced Scorecard-type performance measurement approaches are more common within the R&D community than we had anticipated. Our mail survey of R&D organizations in the private and public sectors, and in academia, suggests that Balanced Scorecard—or some variant of it—is the third most common performance measurement framework among organizations. It accounts for 20% of all respondents, second only to MBO and TQM. It is used more in development labs than in basic research, more in larger labs than in smaller labs, more in public sector and private sector labs than in academic labs, and more in divisional labs than in corporate R&D labs. This also partially confirms our hypothesis about relative prevalence of the Balanced Scorecard (HS1.2). This estimate may be biased downwards as our sample is biased towards academic labs at the expense of public and private sector labs.

The first and foremost principle—or structural feature—of the Balanced Scorecard is its emphasis on performance multi-dimensionality and "balance" across these dimensions. With regards to measurement of multiple attributes of performance—i.e. prevalence of multi-attribute performance measurement systems (MAPMS)—we learnt that their use was quite prevalent among R&D organizations. We used both survey and case study research to assess this feature of R&D organizations. For the purpose of survey research, we used a generic version of an R&D Balanced Scorecard that emphasized on five key dimensions of performance within R&D settings. The survey findings suggest that, on average, R&D organizations considered 3.06 of the five dimensions to be critical to their performance. While this is far from perfect, it is still evidence of the fact that R&D organizations take a multi-dimensional view of performance, even though it may not be "balanced" (a la Kaplan and Norton). This is a confirmation of our prior hypothesis (HS1.3) on the subject.

Case study evidence seems to support the above conclusion. Of the six case study organizations selected for the purpose, three explicitly used a Balanced Scorecard-type performance measurement architecture while the rest also employed a multi-attribute performance measurement system that explicitly defined multiple-dimensions of performance and measured it. The survey results also partially support our prior hypothesis on relative propensities across sectors to view performance as a multi-dimensional construct (HS1.4). Public sector labs, on average, take a more multi-dimensional view of performance than the private sector or academia. These results, however, are not statistically significant.

8.1.2—Balance, participation, accessibility and transparency of strategy-making in R&D—A "differentiated" view of strategy, as advocated by Porter (1996), is the centerpiece of the Balanced Scorecard's strategic-planning architecture. It is hypothesized to provide "laser-like" focus to an organization's efforts to execute its strategy. We used a mix of survey and case study research to

assess various constructs related to strategy-making within the R&D context. The results reveal a mixed picture. For instance, the survey results reveal that R&D organizations take a "diffused" rather than a "differentiated" view of strategy. This is in sharp contrast to the prescription by Kaplan and Norton (2001) that requires organizations implementing a Balanced Scorecard to make choices by adopting one of the several generic strategic themes, with the expectation that a more focused organizational strategy would result in better organizational outcomes.

The results of our survey indicate that R&D organizations, on average, identify 1.72 strategic themes as being applicable to them. Public sector organizations identified 2.24 themes, on average, while academic labs identified 1.3, on average, with private sector labs scoring around 1.70. These results for public-sector and academic labs are different from the private sector labs at a statistically significant level of 5 and 10% respectively. Moreover, better performing labs (i.e. those describing themselves as world class or among the top-quartile) had, on average, a greater number of simultaneous strategic themes than the low-performing labs (i.e. those that identified themselves as above-average, or average). This result is also statistically significant at 5% level.

These results are also confirmed by the case study and interview evidence where we failed to find the kind of strategic sophistication that would support a "differentiated" view of strategy among our population of interest. While we discuss several alternate interpretations of these findings, if true, these could potentially lend some credence to the notion that R&D organizations are different in some fundamental way from other types of commercial organizations and would require that the generic strategy template of the Balanced Scorecard be modified to incorporate this reality of R&D organizations. This finding also fails to confirm our hypothesis (HS2.1) that R&D organizations take a differential view of strategy.

The case study evidence also sheds considerable light on the process dimension of organizational strategy—as it applies to R&D settings. Contrary to our expectations and the common understanding within the R&D community, we found a general trend towards greater employee participation in the strategy-making process—not only within private sector labs but also in public sector and academia. To be fair, while a few seem to get it right, the majority of organizations studied and interviewed are learning to walk the ropes of this rather tricky and "soft" area of organizational management and make it work to their advantage. The overall trend, nonetheless, is unmistakably in the direction of a greater rather than lesser involvement of employees in organizational strategy-making. While there is some evidence to support a trend towards greater satisfaction among organizations implementing more participative and transparent strategy-making processes, we found little conclusive evidence in support of better organizational outcomes. Here, our attempts to identify the effects of participative strategy-making were hindered by lack of temporal maturity of these efforts (HS2.2).

We also looked at the explicit use of performance multi-dimensionality in the strategy-making process itself. The Balanced Scorecard achieves this through the use of explicit cause-and-effect models of performance during the strategy-making exercise. The case study results informed this aspect of the strategy-making processes among R&D organizations. We found R&D organizations seriously lacking in this respect. While several of our case study subjects emphasized multiple dimensions of performance, this formed only one of the several sources of strategic guidance that went into the strategy-making process (HS2.3). For the majority of our case study subjects, explicit recognition of performance multi-dimensionality in strategy-making appeared to be an afterthought rather than the driving factor. While we did find some evidence of strategic learning, albeit at a more limited scale than prescribed by the Balanced Scorecard literature, we also found several pieces of evidence to support that our case study subjects did not realize several of the hypothesized benefits of participative strategy-making, e.g. communication and employee awareness of organizational strategy (HS2.4).

8.1.3—Use of "balanced" scorecards and multi-attribute performance architectures in R&D—The use of "balanced" scorecards that rely on metrics from several dimensions of organizational performance is a norm, rather than an exception, among our survey respondents and case study participants alike. A vast majority of our survey respondents measured multiple dimensions of performance, as did each of our six case study participants (HS3.1). While these above findings are encouraging there are significant omissions as well and hence considerable room for improvements. For example: About 30% of R&D labs studied do not consider employee morale and creativity, 35% do not consider customer satisfaction, and about 40% do not consider innovation management as critical dimensions of performance. Consequently we found some discrepancies in the manner in which performance multi-dimensionality featured into the strategic management systems of our case study subjects e.g. several organizations tended to reward and incentivize performance along key performance dimensions without actually having the systems in place to measure it.

Both as a cause and an effect of the above omissions, therefore, our case subjects also, quite uniformly, did not use explicit cause-and-effect models of performance to "measure the strategy" (HS3.2). This, we believe, was the most glaring of all omissions that seemed to have the most serious effects on the structure and outcomes of Balanced Scorecard-type performance measurement systems. Clearly, several of our case study subjects had either not given enough thought to or had failed to keep up with the evolution of the process aspects of the Balanced Scorecard with the result that had left their implementation of the system seriously compromised. None of the organizations studied managed to realize the kinds of performance improvements hypothesized by the founders of the Balanced Scorecard. While one may only hypothesize about the impact of the counterfactual (i.e. what would have happened had these R&D organizations used a technically sound model of the Balanced Scorecard), and we do so in our discussion on the findings from survey and case study research in their respective chapters, the current research, for lack of identification, cannot claim to

comprehensively speak to and attest for, with appropriate evidence, the performance impact of using a Balanced Scorecard (HS3.3).

8.1.4—Use of incentives to align individual motivations with organizational goals—The apparent consensus on the use of incentives to influence individuals' performance in R&D seemed to be captured in the following phrase by Joseph Stiglitz: "Incentives matter, it matters what type of incentives" (Stiglitz, 2003). This is true for R&D as much as it is true for other areas of economic activity. While there seemed to be considerable philosophical debate on the efficacy of using incentives, especially high-powered financial incentives, in influencing performance, the evidence on the ground was overwhelmingly supportive of the fact that organizations do seem to believe in the power of incentives to influence performance (HS4.1). Both survey and case study evidence suggests that R&D organizations use incentives—financial as well as professional—to align individual and organizational interests.

The propensity to use one type of incentive over the others varies across sectoral boundaries, primarily because of type of work performed, and bureaucratic and equity concerns. Private sector labs, for example, trade individual autonomy and flexibility for a stronger (more intense) incentives regime while academic/university labs do the reverse trading in financial payout in return for greater autonomy and flexibility to choose ones' own research agenda. Public sector labs—constrained by bureaucratic civil-service regulations, on the one hand, and mission requirements, on the other, tend to compensate for their inability to provide high-powered financial incentives by providing the flexibility of picking one's own incentives (a "menu of incentives" approach). We, therefore, failed to confirm the hypothesis on sectoral stereotypes of incentives environment (HS4.2) in its present form.

While we do find evidence suggesting that prevalence of financial incentives follow the pattern described in HS4.2, it is certainly not true if one accounts for all types of incentives. The cumulative survey and case study suggests that various organizational types (e.g. public, private, academic labs) pick and choose an incentives regime that may comprise different bundles of monetary and professional incentives but the overall monetized value of these incentives regimes may not necessarily be markedly different from each other. This way of looking at incentives structures in R&D is also consistent with the notion of self-selection by employees in the type of institution that best provides a mix of incentives that meets their fancy.

Most R&D organizations that we studied, however, believed in the ability of the incentives and recognition to inspire people—an ex-post effect—rather than encourage them to work smarter—an ex-ante effect. For the latter, organizations tended to depend upon rigorous upfront selection processes designed to identify and pick high performers whose career objectives and professional interests matched with the organization's goals and objectives. Seeking to arrive at a match between individual motivations and organizational goals and bridging the discrepancies between the two

with incentives that build upon the self-selection logic to mildly nudge individual behavior towards organizational outcomes is a fundamentally different way of thinking about incentives than one that views incentive intensity or the preponderance of financial or professional incentives as a key characteristic of an organizational incentives system. Needless to say, we could not find strong irrefutable evidence in support of either of the sub-propositions of HS4.3.

8.1.5—Alignment in Strategy, Performance, and Incentives Systems—Alignment between strategy-making, performance measurement, and incentives systems is central to realizing the kinds of performance improvements often attributed to an effective Balanced Scorecard implementation. We assessed the alignment between strategy, performance, and incentives systems in two different ways. Using the mail survey, we assessed the "degree of emphasis" organizations paid to each of the five performance dimensions comprising our generic R&D Balanced Scorecard. The underlying logic behind this analysis was that organizations systematically emphasizing upon a particular (or all five) dimension(s) of performance in their strategic planning, performance measurement, and incentives systems should be able to achieve better performance outcomes. The survey results, however, fail to find such an effect between groups of high-performing and low-performing labs in our sample. Although counterintuitive at first sight, we discuss several alternate explanations for this finding.

The case study evidence provides some credence to our hypothesis (HS5.1). Of the case study subjects, those that have attempted to align various elements of their strategy, performance, and incentive systems with each other as well as with other pieces of the larger organizational strategic management system (e.g. intra-organizational communication and budgeting systems) seem to report higher satisfaction from and greater acceptability of these systems among their sub-organizational units and employees. There is also some qualitative evidence to suggest that organizations that have better aligned strategy, performance, and incentives systems tend to do better in organizational outcomes. This finding clearly points towards a need for a more detailed investigation of alignment of organizational management systems in R&D along the lines suggested by the more traditional management literature. It also highlights the need to debunk the tendency of wholesale rejection of management frameworks based on the popular conception of R&D organizations being different and the need for an unbiased approach to evaluating evidence on the efficacy of the individual organizational interventions (as well as the cumulative package) on performance in R&D settings.

The case study evidence also points towards the effect of selective adoption of individual interventions on the performance impact of the overall package. R&D organizations that tend to show a greater depth of commitment to implementing a Balanced Scorecard-type performance measurement architecture seem to do better, both in reported satisfaction as well as impact on organizational outcomes, than those that selectively adopt the more convenient ones. In fact, we find some qualitative evidence to suggest a non-linear relationship between degree of commitment to the

framework, as defined by the "completeness" of its implementation, and the impact it might have on organizational performance outcomes. More specifically, the adoption of the relatively easier and isolated set of organizational interventions (e.g. thinking about performance in a multi-attribute context,) while providing some proverbial "low hanging fruit", do not begin to bear fruit until more comprehensive set of organizational interventions (e.g. using cause-and-effect performance models, and linking the resultant strategy-performance architecture with the larger strategic management system) are adopted in both letter and spirit. This problem is well-documented in the literature on the traditional management literature and R&D organizations seem to be no exception.

8.2—Conclusions & Future Directions

What does the current research tell us about structuring a Balanced Scorecard for R&D organizations? In addition to the specific minutiae of implementing Balanced Scorecard-type strategy and performance architectures, one can encapsulate our findings into a set of broad and general prescriptions for an R&D organization.

First, the notion of performance multi-dimensionality is well-established in R&D literature and practice. While the intellectual development of R&D measurement literature has lagged behind similar contemporary literatures in certain respects, the lack of a clear bottom-line in R&D has long forced practitioners to think about multiple performance dimensions. Whether the debate was pitched as one about making trade-offs between knowledge creation and commercialization, public or private research, or the research and teaching missions in academia, the notion of performance multidimensionality, implicitly or explicitly, has been at the heart of these discussions. While the multi-attribute frameworks used had limited similarity to the performance dimensions emphasized by the Balanced Scorecard, their prevalence and acceptability within the R&D community makes it easier for these organizations to implement the latter. Our results also indicate that Balanced Scorecard-type multi-attribute approaches to performance measurement are not as uncommon among R&D labs as is normally perceived. An R&D organization contemplating implementing a Balanced Scorecard would, therefore, not be in totally uncharted territory. As the results of this study indicate, one can learn considerably from the examples of other organizations that have implemented such approaches in the past.

Second, while there is some truth to the notion that R&D is different, the idea that it may be different enough so as to render the use of the Balanced Scorecard—or similar management approaches—difficult or even impossible to implement is far from reality. On the one hand, the current research provides considerable evidence in support for using transparent and participative strategy-making processes, multi-attribute performance measurement approaches, cause-and-effect performance models, and the ability of incentives, if used intelligently, to align individual motivations with organizational goals and objectives. On the other hand, however, it also brings to light certain features of R&D organizations that make them different from other types of economic

organizations. For example, that R&D organizations take a diffused rather than differentiated view of strategy is one such finding. Looking at several detailed case studies, however, it is evident that one may be able to modify the Balanced Scorecard to reflect the peculiarities of R&D organizations thus making it more acceptable and amenable to these specific audiences. These modifications may range from stylistic and superficial in some instances (e.g. renaming a performance category from "internal business process" to "managing the innovation process") to more substantive in others (e.g. rethinking generic R&D strategic themes for strategy-mapping.)

Third, the process of implementing a strategy-performance architecture is as important as the substance of the architecture itself. Process is one aspect of the management frameworks—like the Balanced Scorecard—that receives little attention by R&D leaders and managers. It is often viewed with scorn as an "overhead" that takes away a lot of time that could have been spent in more productive research pursuits. Yet, our results indicate that the process of implementing a strategy-performance architecture is perhaps as important as the structure itself. This is especially true of various elements of the strategy-making process as it requires not only that all possible strategy options be put on the table and be open to questioning but also that an organizational consensus must emerge on the best possible strategy. It is through this process that an organization's employees reach a shared understanding of its strategy and build a coalition motivated to execute that strategy. The current research highlights the importance of careful and patient attention to the process aspect of implementing strategy-performance systems—as prescribed by the founders of the Balanced Scorecard movement. In some instances, as we found in our case studies, this may even require demonstrating and publicizing initial successes, providing necessary training, and even a second chance to skeptical employees who might be converted to new ways of doing things.

Fourth, selective adoption of certain features of the strategy-performance framework for implementation may have a disproportionate effect on desired results. R&D organizations planning to implement a Balanced Scorecard must pay careful attention to the overall logic of how it works to bring about the promised performance improvements within organizations. Balanced Scorecard, like every other management framework, is a set of organizational interventions that are related to and build upon each other to deliver results. Selectively adopting certain features this interrelated set of organizational interventions may not only seriously undermine the effectiveness of the overall package but also that of the selected feature being implemented itself. We used the typology of structural and derived features to illustrate the various interdependencies between these features.

For example, the ability to communicate the organization's strategy to each and every individual is one of the most important features of the Balanced Scorecard. It is, however, derived from the use of strategy maps which are in turn dependent upon the explicit use of a cause-and-effect performance model of the organization. In the absence of the latter, therefore, it would be illogical to expect the Balanced Scorecard to provide the kind of strategy awareness that it would otherwise provide. It is therefore important that organizations implementing a Balanced Scorecard seriously think about

and be mindful of the impact on the underlying operating logic of the framework before selectively adopting certain features for implementation and to avoid doing so in an arbitrary fashion.

Fifth, a preliminary assessment of the impact of the Balanced Scorecard on organizational outcomes in R&D suggests the need for some re-calibration of the structure of and expectations from such a system. While the study, for lack of temporal maturity and identification, could not convincingly answer the question of efficacy, it did provide some preliminary insights into the sort of recalibration needed to make the Balanced Scorecard work better in R&D settings. That R&D organizations are not different from other forms of economic organizations is true in many respects. However, there are significant differences as well. Balanced Scorecard-type strategy-performance architectures may need to be recalibrated and modified to better reflect the different reality of R&D organizations. One such difference, as alluded to earlier, is the diffused vs. differentiated view of strategy. Another difference is the manner and extent to which different types of R&D professionals (e.g. scientists vs. engineers vs. technicians) respond to different types of incentives. Yet another difference is the relatively long-term nature of organizational strategy and the loose performance feedback loop that is available to inform an R&D organization's strategic directions.

All of these factors need to be taken into account in defining the expectations from such frameworks and the exact operational logic through which these expectations would be realized. In the absence of enough practitioner-experiential evidence, it seems like a challenge to begin to answer some of these questions. Any prescriptions of how the Balanced Scorecard maybe modified may be partially, if not entirely, speculative. This practitioner-experiential evidence, we believe, would automatically be developed as calls for implementing Balanced Scorecard-type approaches become more frequent and vociferous and R&D leaders attempt to implement and modify R&D Balanced Scorecard. Loch and Tapper (2002) documents one such example. This research provides a detailed look at framework itself—its underlying structural and operating logic, as seen from the lens of an R&D practitioner—and represents a starting point for how R&D organizations may think about and modify it.

Beyond these specific recommendations, the current research also sheds light on several philosophical issues that require a closer examination and some potential directions for future research.

First, that organizational practice often precedes academic theory is an oft-mentioned regularity in traditional literature (e.g. Mohrman and Mohrman, 2004.) The idea sometimes manifests itself as "technology preceding science"(e.g. steam engines came before the laws of thermodynamics were developed and understood.) Theoretists and practitioners often operate in different yet related worlds. They also deal with different experiential realities. More often than not, it is argued, practitioners within the field, while responding to actual organizational problems and challenges, end up in previously uncharted territory "doing and making things happen" while the theoreticians

often scramble to keep pace with explaining "why things happen". Theoretists are often both limited in their ability to accurately model the complex and varied organizational constructs that may help explain, theoretically, what practitioners already intuitively know to be true as well as be restricted by lack of actual data on causes and effects until long after the practitioners have already moved far beyond the confines of existing theory. At this point in the development of an idea, the practitioners have often gone far enough that they do not understand several pieces of their experiential reality thus reducing their ability to fine-tune the organizational interventions. By this time, however, enough data has become available for the theorists to begin theory-building and analysis thus providing some of the answers to practitioners' questions and spurring another round of innovations and refinements.

The development and evolution of the Balanced Scorecard in R&D also partly follows that trajectory of intellectual development. Much of what we know today as the Balanced Scorecard is a result of practitioner-driven innovation and refinement (Kaplan and Norton, 2001.) While it is true that this research provides a starting point to substantively think about using Balanced Scorecard-type frameworks in R&D, as did Kaplan and Norton's basic insight about performance multi-dimensionality (Kaplan and Norton, 1992) that jumpstarted the whole movement, it is dependent on practitioner-experiential data to take it forward to the next level of innovation and refinement. This would require considerable effort in documenting organizational case studies, conducting methodologically sound cross-case analyses, and other empirical research aimed at establishing the effectiveness of these organizational interventions under varying circumstances.

Second, and related to the above is the notion of assessing, in a quantifiable manner, the quality of an organization's implementation of a management framework. Hackman and Wageman (1995) allude to this problem as they ponder over the difficulties of determining whether or not a package of organizational interventions, like the Balanced Scorecard, has actually been implemented. Zbaracki (1998) takes the idea a step further by introducing the notion of technical vs. rhetorical implementations. These authors, especially Hackman and Wageman (1995,) differentiate between a set of features that form the conceptual core of a compound organizational intervention and the add-ons that might provide additional benefits but are not germane to the efficacy of the core. Zbaracki (1998) refers to the core as the technical part of the intervention while everything else as the rhetorical part. It highlights the need for not only isolating the technical from the rhetorical parts but also the technical from the rhetorical motivations for implementing the framework itself.. According to this view, an organizational intervention undertaken for the wrong (rhetorical) reasons may not lead to the kinds of benefits expected of one taken for the right (technical) reasons.

The current research advances these ideas through the use of a "check list" of structural and derived features of a Balanced Scorecard. Taking the analogy a little further, the structural features form the conceptual core of the Balanced Scorecard while the derived features are the additional or related benefits that may arise, separately or through the interplay of various structural features. We use

this checklist of structural and derived features to assess the quality and completeness of an instance of Balanced Scorecard implementation in the case study format. This is clearly a start for such an undertaking and must be built upon before we can truly quantify the quality of the implementation of such organizational management frameworks.

Third, one significant finding of the current study of the study was the notion of differentiated vs. diffused strategy. This, if true, is in contrast to the underlying operating philosophy of the Balanced Scorecard and may potentially be the most damaging evidence in support of the notion that "R&D is different" in ways that fails to satisfy the basic assumptions underlying the Balanced Scorecard. Do R&D organizations take a different view of organizational strategy than other forms of economic organization? While there is some anecdotal evidence that might lead to suggestions that it might be so, no one yet has, to our knowledge, analyzed this particular feature of R&D organizations in any convincing manner. This lack of resolution leaves open the possibility for further research in this area. There is a need to investigate the importance (or lack thereof) of strategic postures in R&D settings, develop a taxonomy of finite (small) number of generic strategic postures that might explain a variety of different R&D organization types and mission requirements, and integrate these with the Balanced Scorecard framework for R&D labs.

One potentially fruitful starting point for this analysis may be Crow and Bozeman's (1998) 3x3 environmental taxonomy of organizational types in R&D that explicitly incorporates the deficiencies of the three-part stereotype of R&D labs (i.e. basic, applied, development) or some other variant of such a scheme. The key underlying factor in the success of such an undertaking would be a keen understanding of the strategic planning process and cognitive frameworks used by R&D leaders and managers in relation with the organizational and mission profiles of their labs. Alternatively, or subsequently, this exercise might also involve informing and stretching the cognitive models of researchers and research managers by helping them become more comfortable with a "differentiated" view of strategy. This would entail pushing a differentiated view of strategy within the "range of convenience" of the cognitive processes of scientists and researchers and research managers as suggested by Reger et al. (1998.)

Fourth, the current research also highlights the tremendous methodological difficulties in doing organizational research of this nature, as rightly outlined by Hackman and Wageman (1995.) These range from definitional issues (i.e. what constitutes an implementation of the Balanced Scorecard) to identification (i.e. how to demonstrate efficacy, or lack of it, without having a set of organizations that have actually implemented a Balanced Scorecard in its true spirit) to causality (i.e. how to attribute an improvement in organizational outcomes to the Balanced Scorecard, rather than any other contemporaneous factor.) This research uses multiple methodologies and multi-case designs to address these issues in several ways. We make considerable progress in some of these challenges and not-so-much in others.

For example, we address the definitional concerns by developing a comprehensive and detailed model of the Balanced Scorecard using Kaplan and Norton's version of it as the point of reference. An exhaustive checklist of structural and derived features, thus developed, provides us with a first-order identification of when an organization has implemented a Balanced Scorecard as well as rough estimate of the quality of its implementation. We address the concerns about identification by using a thought experiment that uses varying levels of implementation across several (case study) organizations and extrapolating the effects observed to develop a sense of which features of the Balanced Scorecard seem to work and which do not as well as how they might all work together. While this is clearly not perfect, it is, nonetheless, the best possible approach to analyzing and drawing conclusions from the evidence at our disposal.

We acknowledge the possibility of encountering non-linear effects when piecing different features of the Balanced Scorecard together and the need to incorporate diversity of organizational forms and missions when extrapolating these findings to other situations. This research highlights the dire need for further work aimed at improving our ability to identify effects of complex organizational interventions but also attributing causal interpretations. Extensive research using natural experiments and pre-post designs in organizations could go a long way in improving our understanding of the impact of these organizational interventions on various constructs of interest. That establishing causality in case study or cross-sectional surveys is tricky, however, should not discourage researchers from the important task of assessing the effectiveness of the Balanced Scorecard in improving organizational outcomes. There is a definite need to conduct thorough and unbiased assessment of Balanced Scorecard implementations.

Finally, the current research also highlights the potential benefits of research designs that employ simultaneous use of multiple methodological approaches. We have already alluded to some of these before, however, this is an important point that deserves a concluding thought here. The decision to use multiple methodological approaches for this research was made to cater to the diverse set of questions needed to address a real-life policy problem. We juggled with the fact that we were dealing with multiple issues (e.g. the whys and the hows) that lent themselves to different types of methodological approaches. The survey component of the research was conceived to answer questions of prevalence and provide us with a sense of the relevant universe. The case study component was designed to analyze Balanced Scorecard-type organizational management interventions in their rich contextual details and develop a sense of their usefulness and efficacy. As the research progressed, however, we found certain additional benefits that, although unexpected, had a major influence on our findings.

These unexpected benefits arose from the use of survey and case study methods in an explanatory and corroboratory manner. By putting the evidence gleaned from one of these approaches to a test against the other approach, we were able to greatly enhance the validity of our findings and discover new ways of interpreting our findings that would not have been possible otherwise. For

example, while the survey research indicated that a certain proportion of R&D organizations used the Balanced Scorecard, it was only through corroboration with the case study evidence that we discovered how different organizations might interpret the term Balanced Scorecard as meaning different things. Similarly, the survey research indicated that R&D organizations used different types of incentives systems to meet their organizational profiles, a finding that was confirmed through corroboration with case study evidence. Thoughtful and careful use of multiple methodological approaches, we believe, has the potential to allow the researchers to go beyond what would be possible within the limitations of a single methodology and, in the process, make possible the analysis of highly complex and contextual organizational issues like the one under consideration. This research clearly demonstrates the potential and pitfalls of that process and is a step forward in that direction.

APPENDIX A—SURVEY QUESTIONNAIRE & COVER LETTER

The Vice Provost For Research/ Chief Scientific Officer/Director of Research,
XYZ University/Lab,
ABC St.,
IJK City, State Zip

Subject: **Your Participation in a Benchmarking Study on Performance Measurement Systems & Incentive Design For Research and Development (R&D) Organizations**

Dear Colleague:

An often-cited factoid in the R&D management literature is that for two years (1993 & 1995) R&D managers responding to the Industrial Research Institute's annual survey rated R&D Measurement as the top ranked problem that they face. What this means is that for a manager of the R&D process, like you, performance measurement still remains somewhat of a challenge. Each one of us would like to have a better idea of the overall health of our R&D organization to be able to not only justify its virtues to external stakeholders but also to use accurate and relevant information to fine tune performance and foresee and navigate potential challenges. You are certainly not alone!

RAND is undertaking a **benchmarking study of 1000 top R&D performers in the US**—industrial corporations, university labs, and federal/defense labs—to better understand R&D measurement and incentives systems in the organizations with varying contexts. Conceived as a project to help a defense client in the midst of a major reorganization, this study has the potential to help inform the performance and incentives debate at both empirical and conceptual levels. Most importantly, we hope to gain insights concerning the relationship between performance measurement, incentives and employee performance so as to help us **develop better quality multi-attribute performance measurement systems and ways to align incentives with organizational performance—** i.e. help yourself and others in the R&D management community learn from US-wide best practices and do a better job at it.

The survey is entirely optional, and can be sent in anonymously. ***We would strongly encourage you to take a few minutes to answer the attached questionnaire for your organization's entire R&D operation and/or copy and forward it to the directors of major R&D labs (corporate, business, or departmental) in your organization***. The data would be kept confidential and analyzed in aggregates only. As a sign of our gratitude, we would be honored to share an advance copy of our research findings with you.

Sincerely,

Bruce J. Held, J.D.
Study Director

Athar Osama
Principal Investigator

Study Advisors:
Steven W. Popper, Ph.D.
Parry M. Norling, Ph.D.
Richard Hillestad, Ph.D.

"PERFORMANCE MEASUREMENT SYSTEMS & INCENTIVE DESIGN FOR RESEARCH AND DEVELOPMENT (R&D) ORGANIZATIONS"

A RAND BENCHMARKING STUDY

INFORMATION FOR THE RESPONDENTS:

1. **Optimal unit of analysis** would be corporate or university R&D labs, research centers, large research groups or the overall company R&D function. Only one response for each sub-organization or facility is requested. Please feel free to pass a copy of this survey to other R&D units within your organization

2. The most appropriate respondents for the survey are: **Chief Scientific / Technology Officers, Vice Presidents of Technology and/or Research, Research Directors, R&D Lab Directors / Managers or Heads of Units responsible for New Product Development and Innovation.**
3. Most of the questions are multiple-choice. You are encouraged to reply to as many questions as possible and as completely as possible. When necessary, use Not Applicable (NA) or Don't Know (DK).
4. **A self-addressed and stamped envelope** is enclosed to return the survey.
5. While the survey maybe answered anonymously, respondents may choose to provide contact info to:
   a) **Receive an advance copy of the detailed findings** of this study.
   b) **Participate in a post-survey, in-depth case study that may make possible tailored organization-specific feedback (e.g. Executive Policy Brief) to the responding organization**

Analysis will be conducted on aggregated data. Under no circumstances will the identities of participating organizations be disclosed, except as required by law. The data will only be used for research purposes and findings might be published as a RAND publication and/or in a major publicly available scholarly journal. The study is approved by RAND's Human Subjects Protection Committee.

INFORMATION ABOUT THE CONDUCTING INSTITUTION:

This research is being undertaken at RAND—a private, not-for-profit corporation that conducts objective, fact-based research to better inform decision-making on public policy issues. (Please see: http://www.rand.org for details.) Questions regarding this survey or the subsequent research maybe be directed to:

Contact address:
Bruce Held J.D.
RAND Arroyo Center
Tel: 1-310-393-0411 ext. 6745, Fax: 1-310-451-6978
Email: RDproject@rand.org

Mailing address (for survey):
RAND
Attn: R&D Performance Project
1700 Main St.,
Santa Monica, CA 90407

PART 0: CONTACT INFORMATION (OPTIONAL–Fill Only If applicable, see 6a, 6b above**) STUDY ID:** _____

Name of the Organization: _____

Name of the Sub-organization: _____

Contact Person: _____ Position: _____ Email: _____

Contact Address: _____

_____Tel: _____Fax: _____

Do you wish to consider participation in a post-survey case study? Yes [ ], No [ ] ***(Please provide contact info)***

Do you wish to receive the final report / results (please tick)? Yes [ ], No [ ] ***(Please provide contact info)***

PART 1: CORPORATE / R&D SUB-ORGANIZATION STRUCTURE AND PERFORMANCE INFORMATION

Please answer the following concerning the overall organization or company

1. Which of the following best describes the ownership structure of the organization? *(Please mark one)*

☐ Publicly-held Corp. ☐ Privately-held Corp. ☐ Public Sector / Govt. ☐ University ☐ Other _____

2. How many ***employees*** does the organization have? *(Please mark the most appropriate)*

☐ < 100 ☐ 100 - 499 ☐ 500 - 4999 ☐ 5000- 49,999 ☐ > 50,000

3. Please provide the following information about this company or organization:

a. Is this company a US-subsidiary of a foreign company (please mark) ☐ No, ☐ Yes
b. Your key business or product areas (describe or write 3-digit SIC)
_____
c. Corporate sales or revenues (US$)_____Year: _____
d. Corporate Profit in the above year (US$) _____
e. Corporate R&D expenditure in the above year (US$)_____

4. Of the current product line, what proportion is ***new products or services*** introduced in last 5 years *(Please mark the most appropriate, see page.8 for definition of new products or services)*

☐ > 50% of current products ☐ 34-49% of current products ☐ 11-33% of current products

5. What is the breakdown of R&D expenditure in terms of where it is spent? *(Please specify as %)*

☐ Central lab _____ ☐ Business Units labs _____ ☐ External / Contract / Other _____ = 100%

PART 2: R&D LAB/GROUP/TEAM (UNIT OF ANALYSIS)–STRUCTURE & PERFORMANCE INFORMATION

(The following questions must be answered for the part of the company's overall R&D organization for which this questionnaire is being filled—e.g. the entire company R&D function, a corporate Lab, a divisional or department lab or any other unit of analysis henceforth referred to as "this facility".

6. Please specify the ***organizational entity or "facility"*** for which the questionnaire is being filled.

☐ Overall company R&D ☐ Central Lab ☐ Divisional / Department Lab ☐ Other (specify) _____

7. Please provide the following information about this facility:

a. Most recent estimate for ***annual R&D expenditure*** (US$)_____Year _____
b. Average capital investment (facilities and equipment) in last 3 years: (US$)_____

8. How is ***R&D expenditure*** divided in terms of the type of activity it is spent at this facility *(Please specify as % of total budget, see page.8 for definitions of terms)*

Basic Research _____ Applied Research _____ Product / Process Improvement _____
New Product Development _____ Test & Engineering _____ Technical Support[94] _____ = 100%

9. What is the best estimate of the **# *of technical and managerial employees*** at this facility? *(Please mark the most appropriate)*

[94] For the purpose of this survey, Technical Services is defined as "providing consultations or troubleshooting to clients or business units on a short-term basis."

☐ < 20 ☐ 20 - 100 ☐ 100 - 500 ☐ 500- 1000 ☐ > 1000

10. What is the ***mix of educational qualifications of technical and managerial employees*** at this facility? *(Please specify all percentages)*

PhD/ MD/M.Phil. _____ MSE/MS _____ BS/BSE _____MBA _____ Other _____ >= 100%

11. How does the average salary level of technical professionals at this facility compare with its peers?

☐ Among the top-5 leaders in terms of salaries
☐ Rated consistently as higher-than-average salaries
☐ Equal to or slightly above industry average salaries
☐ Pays below the industry average salaries

12. Estimates of the following ***indicators of R&D performance*** of this facility *(please provide specific #)*

☐ # of new product introductions (last 5 yrs.): _____ ☐ Not Applicable ☐ Don't Know
☐ # of new projects initiated (last 5 years): _____ ☐ Not Applicable ☐ Don't Know
☐ # of significant product improvements (last 5 yrs): _____ ☐ Not Applicable ☐ Don't Know
☐ # of significant process improvements (last 5 yrs): _____ ☐ Not Applicable ☐ Don't Know
☐ # of new patents granted (last 5 yrs): _____ ☐ Not Applicable ☐ Don't Know
☐ # of peer-reviewed papers published (last 5 yrs): _____ ☐ Not Applicable ☐ Don't Know
☐ # of successful project transitions (last 5 yrs): _____ ☐ Not Applicable ☐ Don't Know
☐ Any other significant indicator of performance (last 5 yrs): _____
☐ Any other significant indicator of performance (last 5 yrs): _____

13. How would you rate the overall performance of this facility with respect to its peers within the industry segment (or type of work etc.)? *(Please mark the most appropriate)*

☐ World leader (engaged in cutting-edge R&D)
☐ In the top quartile (~ among the top-5 facilities in class)
☐ Above average
☐ Average or Other (specify) _____

14. Can you identify an R&D lab or research center in your field that has the reputation of being:

a. Best / high performing in its class / sector: _____
b. Most dramatic turnaround in recent years: _____

PART 3: ORGANIZATION'S (OR SUB-ORGANIZATION'S) R&D OR TECHNOLOGY STRATEGY
(This section examines focus and alignment of various elements of the R&D lab's [technology] strategy)

15. How would you characterize the Technology/R&D strategy (or strategic posture) of this facility? *(Please mark as many as appropriate)*

☐ Technical Excellence (Innovative Leadership)
☐ Operational Excellence (On-Time, On-Cost, On-Specs)
☐ Customer Responsiveness (Internal or External)
☐ Other (specify) _____

16. Which of the following do you consider critical dimensions of performance at this facility? *(Please mark as many as appropriate)*

☐ Employee morale and creativity (EMC)
☐ Innovation management (IM)
☐ Financial performance and control (FPC)
☐ Other (specify) _____

☐ Learning and knowledge management (LKM)
☐ Customer satisfaction (CS)
☐ Other (specify) _____
☐ Other (specify) _____

17. Please mark which of the following statements are true for each of the five dimensions of performance identified in the above item (#21) at this facility.

DEGREE OF EMPHASIS PAID TO 5-KEY PERFORMANCE DIMENSIONS	EMPLOYEE (EMC)	LEARNING (LKM)	INNOVATION (IM)	CUSTOMER (CS)	FINANCIAL (FPC)	OTHER
Considered important by management	☐	☐	☐	☐	☐	☐
Management emphasizes it as a source of strategic advantage	☐	☐	☐	☐	☐	☐
Explicit initiatives exist to develop a strategic edge in this area	☐	☐	☐	☐	☐	☐
Performance in the area is regularly measured and monitored	☐	☐	☐	☐	☐	☐
Rewards and incentives are tied to performance in the area	☐	☐	☐	☐	☐	☐

PART 4: THE ORGANIZATION'S PERFORMANCE MEASUREMENT PHILOSOPHY

(This section asks questions about the uses of performance measures)

18. Please indicate your ***degree of agreement*** with the following ***potential uses of measurement systems*** in an R&D setting? *(Please mark each statement from "1" for "Strongly Disagree" to "5" for "Strongly Agree")*

	Strongly Disagree ⟷ *Strongly Agree*					
a) Identify problems/deviations for action by managers	☐ 1	☐	☐	☐	☐ 5	☐ Don't Know
b) Improve employee understanding of the org. systems	☐ 1	☐	☐	☐	☐ 5	☐ Don't Know
c) Align and communicate organizational objectives	☐ 1	☐	☐	☐	☐ 5	☐ Don't Know
d) Provide input for performance-based rewards	☐ 1	☐	☐	☐	☐ 5	☐ Don't Know
e) Identify problems/deviations for action by employees	☐ 1	☐	☐	☐	☐ 5	☐ Don't Know
f) Justify existence, decisions and performance	☐ 1	☐	☐	☐	☐ 5	☐ Don't Know
g) Motivate people by providing feedback	☐ 1	☐	☐	☐	☐ 5	☐ Don't Know

19. Please indicate the ***actual uses*** of performance information at this facility? *(Please mark as many as appropriate)*

☐ To provide a basis for incentives/recognition
☐ To provide feed back to employees/management
☐ To communicate with corporate management
☐ To guide employee development policies
☐ To aid in R&D policy/strategy formulation
☐ Other *(specify)* _____

PART 5: CURRENT PERFORMANCE MEASUREMENT SYSTEMS AND TECHNIQUES

(This section addresses the specifics of the performance measurement system in place at the facility)

20. Please state if any of the following ***performance assessment and management frameworks*** is currently being used at this facility? *(please mark all appropriate answers)*

- ☐ Total quality management (TQM)
- ☐ Balanced Scorecard (BSC)
- ☐ Technology Value Pyramid (IRI Model)
- ☐ Management-by-objectives (MBO)
- ☐ Six Sigma-type techniques
- ☐ Other *(specify)* _____

21. For how long has the above performance measurement framework(s) been in place? _____ yrs.

22. In your opinion, what kind of impact does the measurement have on the performance at this facility? *(Please mark most appropriate)*

Very Positive	Slightly Positive	No Impact	Slightly Negative	Very Negative
☐	☐	☐	☐	☐

23. Please describe the ***five most important performance indicators*** (metrics or procedures) used to assess the performance of an ***individual, research team/group, lab and/or overall company R&D function*** *(you may assume an "average" individual, group/team working at this facility.)*

TYPE OF METRIC, PERFORMANCE INDICATOR, OR PROCEDURE	**INDIVIDUAL** *(mark any five)*	**GROUP/TEAM** *(mark any five)*	**THIS FACILITY** *(mark any five)*
Bibliometrics (e.g.# of patent, # of papers etc.)	☐	☐	☐
Impact metrics (e.g. # of paper or patent citations)	☐	☐	☐
Return on Investment (e.g. Yield, Net-NPV etc.)	☐	☐	☐
Peer Assessment (e.g. external/internal reviews)	☐	☐	☐
New products/services introduced in last x-years	☐	☐	☐
Improvements in individual/aggregate capability	☐	☐	☐
Time to market metrics (e.g. development cycle)	☐	☐	☐
# of successful/unsuccessful technology transitions	☐	☐	☐
Quality of strategic alignment or goal attainment	☐	☐	☐
Cost over-runs (e.g. # of projects within-budget)	☐	☐	☐
Schedule over-runs (e.g. # of projects delayed)	☐	☐	☐
Measures of teaching/training activities	☐	☐	☐
# or $-value of proposals funded/contract obtained	☐	☐	☐
Professional peer recognition (awards or ranking)	☐	☐	☐
Ratio of direct labor to indirect overheads costs	☐	☐	☐
Revenue/cost savings from new products/services	☐	☐	☐
Self-described R&D effectiveness scores	☐	☐	☐
Customer satisfaction measures	☐	☐	☐
Customer buy-in (e.g. optional business-unit funds)	☐	☐	☐

% of development pipeline milestones achieved	☐	☐	☐
Budgetary growth achieved	☐	☐	☐
Employee morale and creativity measures	☐	☐	☐
% of Projects that meet performance specifications	☐	☐	☐
Efficiency of processes (e.g. success ratio of projects)	☐	☐	☐
Technical competence (e.g. # of PhDs or Post-docs)	☐	☐	☐
Other (specify):	☐	☐	☐
Other (specify):	☐	☐	☐
Other (specify):	☐	☐	☐

PART 6: CURRENT PERFORMANCE MEASUREMENT SYSTEM – INCENTIVES / COMPENSATION LINKAGE

(This section addresses the link between performance measurement and incentive/reward in place at the facility)

24. For what ***type of workforce*** at this facility, are the following ***rules used for making compensation decisions***? *(Please mark the as many compensation schemes as appropriate and mark one of the three boxed choices for type of employee)*

TYPE OF COMPENSATION SCHEME	**MOST EMPLOYEES**	**MANAGERIAL EMPLOYEES ONLY**	**KEY EMPLOYEES ONLY**
☐ Across-the-board cost of living-type % increase	☐	☐	☐
☐ Years-in-service-based pay increase	☐	☐	☐
☐ Individual performance-based fixed increase	☐	☐	☐
☐ Individual performance-based variable increase	☐	☐	☐
☐ Group performance-based fixed increase	☐	☐	☐
☐ Group performance-based variable increase	☐	☐	☐
☐ Other (specify) _____	☐	☐	☐
☐ Other (specify) _____	☐	☐	☐

25. Please specify the type(s) of ***individual-level rewards/recognition/benefits*** that are available (are used or can be expected) used by employees in response to good performance at this facility. *(Please mark as many as appropriate)*

☐ Lab-wide recognition and/or token reward
☐ Corporate-wide recognition and/or token reward
☐ Cash rewards (please indicate the maximum size of cash reward _____)
☐ Creation of new strategic business units (spin-offs) or entrepreneurial sabbaticals
☐ Fixed cash reward per patent/invention/paper etc.
☐ Time/money to attend professional events, conferences, skill-development activities
☐ Share of royalties on patents/inventions, please state if capped or uncapped: _____
☐ Access to discretionary funding (or time) to pursue interests or publications

☐ Improved ability to work on prestigious projects the following year
☐ Better independence and autonomy to pursue personal research agenda
☐ Chances of more frequent (or fast-track) promotion within the technical track within the organization
☐ Chances of being considered for a research management position within the organization
☐ Group-based fixed or variable performance-rewards
☐ Greater visibility and recognition among peers, frequency of being invited to "important" meetings
☐ Other (specify)_____

26. Please specify the type(s) of ***laboratory-level benefits/ rewards / recognition*** that that are available (are used or can be expected as a result of good performance by a particular lab at this facility. (Please mark as many as appropriate)

☐ Corporate-wide recognition/publicity/token-reward for the lab
☐ Chances of promotion for lab's management
☐ Significant financial incentives for lab's management
☐ Visibility, recognition, chances of promotion for lab's employees
☐ Ability to negotiate better (or more interesting) projects for future years
☐ Ability to secure more funding for the future years
☐ Ability to attract talent from within the organization
☐ Ability of the lab's management to adopt attractive compensation systems
☐ Access by lab's management to discretionary funding (e.g. for projects, incentives, training etc.)

27 Please indicate the following about the ***most significant financial awards*** available to individuals and teams/groups at this facility

FEATURES OF FINANCIAL AWARDS	INDIVIDUALS ☐ If not applicable	TEAMS/R&D GROUPS ☐ If not applicable
The maximum size of each award ($)		
Total # of awards made per year (average of exact #)		
% of individuals eligible for the award [95]		
Is it based on subjective assessment or objective target(s)		

28. Please indicate how the ***most significant non-financial or professional award*** is given to individuals at this facility.

a. What is the most significant professional award called: _____
b. The professional recognition associated with award is at the level of:
☐ Lab-wide ☐ Corporate-wide ☐ Professional community-wide
c. The number of professional awards made each year (average or exact #): _____
d. % of individuals eligible for the award at a given time (please see note below)[2]: _____

[95] This percentage should identify the individuals who can, at a given time, feasibly compete for the stated financial award. For example, the answer would range between 100% (if everyone can feasibly win) to 5% (if only senior most staff does.)

29. Do you have a story to tell about significant or dramatic improvement or decline in organizational performance resulting from the implementation of a performance measurement system or incentives plan? Please share your experiences with us. *(Please use a additional blank page, if you so desire)*

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Please feel free to contact the study director at RDproject@rand.org for clarifications on the survey, and comments or suggestions about the study. Thank you for your kind participation and hope to get back to you with study results – Study Director

APPENDIX B: INFORMED CONSENT FORM

CONSENT FORM FOR PARTICIPATION IN "PERFORMANCE & INCENTIVES IN R&D" STUDY

We request your consent to participate in the case studies component of the above research study.

PURPOSE OF THE STUDY

RAND, a not-for-profit and non-partisan research institution in Santa Monica, California, that aims to improve decision-making through research and analysis is conducting a study of Performance Measurement and Incentive Systems in Research and Development (R&D) Organizations. The study is funded as a part of a contract with the United States Army under the RAND Arroyo Center. The goal of the study is to look at how R&D organizations integrate their strategy formulation, performance measurement and incentive design functions to improve R&D performance. The study would meticulously explore the "possibility" of developing or adapting a multi-attribute performance measurement framework that can be applied in the context of an R&D organization. While such frameworks—especially the Balanced Scorecard—are becoming increasingly popular in business world,

they have been of limited use within the R&D community (e.g. IRI's Technology Value Pyramid). The study would help identify implementation and perceptual bottlenecks that can be the causes for the above phenomenon, try to test the appropriateness of various assumptions behind the Balanced Scorecard framework in the R&D environment, and make recommendations for developing such a framework for R&D.

HOW WE SELECTED YOU

We selected you for possible participation in the study because your organization provides us with an opportunity to study a performance and incentives in an interesting setting while at the same time help us include as many diverse organizations (e.g. public vs. private, basic vs. applied, and academic vs. commercial) as possible. While none of the R&D organizations that we know of has implemented a Balanced Scorecard Approach, by studying varying models of performance measurement systems, we hope to gain valuable insight and test various assumptions of the Balanced Scorecard framework. This would allow us to tailor our model to better suit the R&D environment.

WHAT WE WILL ASK YOU TO DO

Participation in the study might include the following activities:

a) ***Filling out a detailed survey questionnaire*** to provide us with a better understanding of the strategy formulation, performance measurement and incentive design functions at your organization.
b) ***Sharing of organization-specific information*** about systems, structures and processes in place that deal with strategy formulation, performance measurement and incentive design functions. You may choose to mask any proprietary data and share the information in generalities only OR you may choose to share such data for to allow the researcher to better understand the context and impact of the above systems and processes (any proprietary data would nevertheless be kept strictly confidential).
c) ***Interviews with key individuals*** in the areas of strategy formulation, performance measurement, incentive design, implementation and/or bench-level scientists and engineers mutually agreed upon with the lead researcher.
d) ***A field visit to your facility*** (to be conducted around the interview process) to allow the lead researcher to have a better feel of the research environment and meet with a few scientists and engineers as well as provide a face-to-face opportunity to seek more information.
e) ***General information sharing/liaison*** to allow the lead researcher to adequately document the results of the case study, seek clarifications for ambiguous or missing information.
f) ***Review and comment on the draft copy*** of the individual case study furnished to you within one month of the start of the field research process.

RISKS OF PARTICIPATION & STEPS TOWARDS MITIGATION

The study poses minimal foreseeable risks for participants. Among the possibilities include: minor personal or professional consequences for individuals sharing information that might be perceived as negative by organizational leadership; the possibility of the case study findings being perceived as a cause of harm to organization's reputation among peers; and the possibility of inadvertently publishing information deemed proprietary by the organizational leadership. ***Several steps have been taken to address those risk factors.***

Proprietary Information:
First, the study makes it optional for organizations to share proprietary data with the researcher. However, if they choose to do so to allow the researcher to better understand the organizational context, the data would be kept under strict confidentiality, will not be cited directly, and destroyed immediately after the completion of the study.

Individual Identities:
Second, names of individuals participating in the research would not be cited directly to shield them from any possible personal or professional consequences that might arise from their participating in the study. Quotations, where used, would be kept anonymous. The draft individual case studies would be shared with respective organizations to allow them to comment upon, clarify and/or register dissent to any part of the individual case report. These discrepancies would be sorted out through mutual agreement and where required these comments would be published alongside the case report.

Organizational Identity in Individual& Cross-case Comparison:
While it is the desire of the researcher to use real names of organizations in the individual case studies to improve the overall quality of the work and readers' comprehension of organizational context, identities of participating organizations maybe masked if so requested. Even if the organizations chooses to allow the researchers to use their real identities in the individual case reports, the cross-case analysis would only cite anonymous references to case study participants.

These steps, we believe, would help minimize the risks to the participants of the study.

BENEFITS OF PARTICIPATION

There are several benefits of participation in this research, some accruing directly to the participant and others more dispersed in the form of better understanding of a complex but important issue. Among the direct benefits that might accrue to the participants of the study would be an advance copy of the study findings, as and when they become available. In addition to that, participation in the case study would allow the possibility of tailored feedback to the participating organizations on how they are doing vis-à-vis their peers. This might take the form of the continuous exchange of ideas and feedback during the case study process itself and/or a tailored executive brief to the R&D management of the organization providing a status assessment and possible recommendations and/or an oral briefing, if possible. In addition to these direct benefits, the participating organizations as well as the R&D management community more generally would benefit from an improved understanding of the issue of strategy-performance-incentive alignment. Based upon the insights gained during this research, the final analysis would include a recommendation on whether or not a multi-attribute performance measurement system makes sense in the R&D setting, what form should it take, and what kinds of implementation bottlenecks can be expected if an organization were to decide upon implementing one.

CONFIDENTIALITY

We will use the information you give us for research purposes only. We will protect the confidentiality of this information, and will not disclose your identity or information that identifies you to anyone outside of the research project, except as required by law. We will not identify you personally in any reports we write. The organization's real identity will be used in the individual case study report only if you choose to allow us to do so and that authority can be rescinded at any time leading up to the receipt of your comments on the draft individual case study report. Any information that you identify as proprietary but choose to share with us would not be directly cited in the case study report and would be kept strictly

confidential. We will destroy all proprietary information as soon as it is no longer needed and all information that identifies you at the end of the study.

VOLUNTARINESS

Your participation in the study is completely voluntary. You may refuse to participate, or you may stop participating at any time and for any reason, without any penalty. We may also discontinue your participation or stop the study at any time if circumstances warrant.

WHOM TO CONTACT

If you have any questions about the study, please contact Athar Osama, 1700 Main St. Santa Monica, CA 90407, 310-393-0411 x6745

If you have any questions or concerns about your rights as a research subject, please contact the Human Subjects Protection Committee at RAND, 1700 Main Street, Santa Monica, CA 90407, 310-393-0411, ext. 6369.

CONSENT TO PARTICIPATE

I have read this statement, and I understand what it says. I agree to participate in this study under the conditions outlined above. I also acknowledge that I have received a copy of this form.

Signature_____ Date_____

Printed Name_____ Organization _____

APPENDIX C: SOME RELEVANT STATISTICAL TABLES

TABLE C-1: DIFFERENCES BETWEEN RESEARCH LABS AND DEVELOPMENT LABS ON KEY DIMENSIONS OF RELEVANCE (DIFF OF MEANS TEST)

Important Structural Variables	Aggregate (#,N=85)	Research Labs (#,N=50)	Development Labs (#,N=35)
Self Described Performance Ranking	2.51	2.63	2.34
Sector of Operation Private: Public: Academic	34:30:36	12:32:56	63:28:9
R&D Strategies or Strategic Themes	%	%	%
Technical Excellence	73.8	81**	62
Operational Excellence	25	20	31
Customer Responsiveness	57.1	44*	74
Other R&D Strategy	13	10	17
Total # of Simultaneous Strategies	1.72	1.60	1.88
Degree of "Balance" in Performance Dimensions	%	%	%
Employee Morale & Creativity	71	76	65
Learning and Knowledge Mgmt	60	70*	45
Innovation Mgmt.	57	58	57
Customer Satisfaction	64	52*	82
Financial Performance/Control	49	46	54
Other Performance Dimension	25	16*	40
Average # of Dimensions	3.29	3.18	3.45
Emphasis on Performance Dimensions	#	#	#
Employee Morale & Creativity	2.32	2.63	1.91
Learning and Knowledge Mgmt	1.90	2.53	1.05
Innovation Mgmt.	2.64	2.65	2.62
Customer Satisfaction	2.07	1.85	2.42
Financial Performance/Control	1.79	1.55	2.11
Other	1.95	2.2	1.78
Performance Measurement Systems in Place	%	%	%
Total Quality Mgmt.	23	28	17
Management by Objectives	41	38	45
Balanced Scorecard	20	14	28
Six-Sigma type Techniques	17	8*	31
Technology-Value Pyramid	1	0	2
Other	17	12	25
Total # of Frameworks Used	1.21	1*	1.51
How Long in place (yrs.)	6.01	6.43	5.57
Perceived Impact	3.98	4.08	3.86
Compensation and Incentives Structures			
Salary (self-described)	1.74	1.65	1.85
# of Simultaneous Compensation Options	1.90	1.94	1.85
# of Simultaneous Individual Incentive Options	6.27	6.58	5.82
# of Simultaneous Lab-level Incentive Options	2.85	3.06	2.57

* indicates a 5% significance level, ** indicates a 10% significance level, Private is the benchmark category

TABLE C-2: DIFFERENCES BETWEEN LARGE & SMALL LABS/R&D OPERATIONS ON KEY DIMENSIONS OF RELEVANCE (DIFF. OF MEANS TEST)

Important Structural Variables	Aggregate (#,N=85)	Large Labs (>100) (#,N=29)	Small Labs (<100) (#,N=56)
Self Described Performance Ranking	2.51	2.42	2.55
Sector of Operation Private: Public: Academic	33:30.5:36.5	17.25:17.25:65.5	41:37.5:21.5
R&D Strategies or Strategic Themes	%	%	%
Technical Excellence	73.8	85**	67
Operational Excellence	25	25	25
Customer Responsiveness	57.1	42**	64
Other R&D Strategy	13	7	16
Total # of Simultaneous Strategies	1.72	1.60	1.77
Degree of "Balance" in Performance Dimensions	%	%	%
Employee Morale & Creativity	71	75	69
Learning and Knowledge Mgmt	60	75*	51
Innovation Mgmt.	57	51	60
Customer Satisfaction	64	44*	75
Financial Performance/Control	49	55	46
Other Performance Dimension	25	13**	32
Average # of Dimensions	3.29	3.17	3.35
Emphasis on Performance Dimensions	#	#	#
Employee Morale & Creativity	2.32	2.61	2.19
Learning and Knowledge Mgmt	1.90	2.57*	1.58
Innovation Mgmt.	2.64	2.38	2.76
Customer Satisfaction	2.07	1.61	2.28
Financial Performance/Control	1.79	2.07	1.66
Other	1.95	3.66**	1.71
Performance Measurement Systems in Place	%	%	%
Total Quality Mgmt.	23	6*	32
Management by Objectives	41	20*	51
Balanced Scorecard	20	24	17
Six-Sigma type Techniques	17	10	21
Technology-Value Pyramid	1	0	1
Other	17	13	19
Total # of Frameworks Used	1.21	.75*	1.44
How Long in place (yrs.)	6.01	5.8	6.09
Perceived Impact	3.98	3.87	4.02
Compensation and Incentives Structures			
Salary (self-described)	1.74	1.57	1.82
# of Simultaneous Compensation Options	1.90	1.79	1.96
# of Simultaneous Individual Incentive Options	6.27	6.72	6.03
# of Simultaneous Lab-level Incentive Options	2.85	2.75	2.91

* indicates a 5% significance level, ** indicates a 10% significance level, Small Lab is the benchmark category

TABLE C-3: DIFFERENCES BETWEEN OVERALL COMPANY R&D, CORPORATE LABS, AND DIVISIONAL LABS ON KEY DIMENSIONS OF RELEVANCE (BENCHMARK: OVERALL CO.)

Important Structural Variables	Aggregate (#,N=78)	Overall Co. (#,N=48)	Corporate (#,N=11)	Divisional (#,N=19)
Self Described Performance Ranking	2.53	2.31	2.72	2.94*
Type of Lab (Research: Development)	56:44	62:38	18:82	63:37
R&D Strategies or Strategic Themes	%	%	%	%
Technical Excellence	76	76	81	73
Operational Excellence	27	27	27	26
Customer Responsiveness	57	46	72*	73*
Other R&D Strategy	13	15	18	5
Total # of Simultaneous Strategies	1.75	1.67	2	1.78
Degree of "Balance" in Performance Dimensions	%	%	%	%
Employee Morale & Creativity	71	72	81	63
Learning and Knowledge Mgmt	60	66	54	47
Innovation Mgmt.	55	62	36	47
Customer Satisfaction	64	47	1*	84
Financial Performance/Control	48	50	54	42
Other Performance Dimension	25	27	45	10
Average # of Dimensions	3.25	3.27	3.72	2.94**
Emphasis on Performance Dimensions	#	#	#	#
Employee Morale & Creativity	2.3	2.31	2.54	2.15
Learning and Knowledge Mgmt	1.88	2.15	1.36	1.52
Innovation Mgmt.	2.54	2.73	2.45	2.15
Customer Satisfaction	2.06	1.91	2.45	2.21
Financial Performance/Control	1.81	1.86	2.36	1.36
Other	1.7	1.61	1.6	2.5
Performance Measurement Systems in Place	%	%	%	%
Total Quality Mgmt.	24	20	27	31
Management by Objectives	39	45	27	31
Balanced Scorecard	19	16	9	31
Six-Sigma type Techniques	17	18	9	21
Technology-Value Pyramid	1	2	0	0
Other	19	20	27	10
Total # of Frameworks Used	1.21	1.25	1	1.26
How Long in place (yrs.)	6.01	6.1	5.12	6.30
Perceived Impact	4.03	3.97	4.0	4.21
Compensation & Incentives Structures	#	#	#	#
Salary (self-described)	1.73	1.73	1.54	1.83
# of Simultaneous Compensation Options	1.87	1.83	1.72	2.07
# of Simultaneous Individual Incentive Options	6.21	6.43	6.18	5.68
# of Simultaneous Lab-level Incentive Options	2.76	2.79	2.90	2.63

* indicates a 5% significance level, ** indicates a 10% significance level, Overall Company is the benchmark category

TABLE C-4: DIFFERENCES ACROSS WORLD CLASS, TOP-QUARTILE, ABOVE-AVERAGE AND AVERAGE R&D LABS ON KEY DIMENSIONS OF RELEVANCE (SIMPLE AVERAGES)

Important Structural Variables	Aggregate (N=~84)	WorldC (N=~24)	Top-Q (N=~14)	Ave+ (N=~27)	Ave. (N=~19)
Self Described Performance Ranking	2.51	4	3	2	1
Sector of Operation Private: Public: Academic	33:30.5:36.5				
R&D Strategies or Strategic Themes	%	%	%		
Technical Excellence	73.8	87	92	66	52
Operational Excellence	25	37	14	25	15
Customer Responsiveness	571	66	57	55	42
Other R&D Strategy	13	16	7	11	17
Total # of Simultaneous Strategies	1.72	2.08	1.71	1.62	1.35
Degree of "Balance" in Performance Dimensions	%				
Employee Morale & Creativity	71	79	71	70	63
Learning and Knowledge Mgmt	60	54	50	74	52
Innovation Mgmt.	57	58	71	59	47
Customer Satisfaction	64	71	78	62	52
Financial Performance/Control	49	62	21	66	31
Other Performance Dimension	25	21	28	29	26
Average # of Dimensions	3.29	3.4	3.21	3.62	2.73
Emphasis on Performance Dimensions	#				
Employee Morale & Creativity	2.32	2.4	2.21	2.85	1.44
Learning and Knowledge Mgmt	1.90	2.2	1.71	1.96	1.55
Innovation Mgmt.	2.64	3.1	3.07	2.70	1.55
Customer Satisfaction	2.07	2.8	1.85	2.33	.94
Financial Performance/Control	1.79	2.1	1.35	2.25	1.05
Other	1.95	.8	3.2	1.75	2.6
Performance Measurement Systems in Place	%				
Total Quality Mgmt.	23	29	42	18	10
Management by Objectives	41	45	42	44	31
Balanced Scorecard	20	21	28	25	5
Six-Sigma type Techniques	17	10	35	18	15
Technology-Value Pyramid	1	0	0	3	0
Other	17	.21	.28	.14	.10
Total # of Frameworks Used	1.21	1.25	1.78	1.25	.73
How Long in place (yrs.)	6.01	6.88	4.78	6.88	4.55
Perceived Impact	3.98	4.38	3.85	3.9	3.16
Compensation and Incentives Structures	#				
Salary (self-described)	1.74	1.91	1.85	1.69	1.5
# of Simultaneous Compensation Options	1.90	2.04	2.07	1.81	1.73
# of Simultaneous Individual Incentive Options	6.27	6.37	7.21	7.07	4.47
# of Simultaneous Lab-level Incentive Options	2.85	2.5	3.0	3.33	2.63

BIBLIOGRAPHICAL REFERENCES

Adams, C., and Roberts, P., "You Are What You Measure", In ***Manufacturing Europe 1993***, Sterling Publications, pp. 504-507, 1993

Agarwal, Naresh, and Singh, Parbudyal, "Organizational Rewards for a Changing Workplace: An Examination of Theory and Practice", ***In International Journal of Technology Management***, v.16, n.1/2/3, 1998, pp. 225-237

Aiman-Smith, Lynda, Markham, Stephen K., "What You Should Know About Using Surveys", In ***Research-Technology Management***, 2004, pp. 12-15

Amabile, Teresa M., Conti, Regina, Coon, Heather, Lazebny, Jeffrey, Herron, Michael, "Assessing the Work Environment for Creativity", In ***The Academy of Management Journal***, v.29, n.5, Oct 1996, pp. 1154-1184

Amram, M., and Kalutilaka, N., Real Options: Managing Strategic Investments in an Uncertain World, Boston MA: Harvard Business School Press, 1999

Anderson, Richard E., and Meyerson, Joel W. (Eds.), Productivity and Higher Education: Improving the Effectiveness of Faculty, Facilities, and Financial Resources, Peterson Guides Inc., 1992

AQPC, "NCR Case Study", In ***Performance Measurement: Implementing the Balanced Scorecard***, American Quality and Productivity Council, 2002, pp. 42-59

Argyris, Chris, "Teaching Smart People How to Learn", In ***Harvard Business Review***, Boston: Harvard Business School Press, May-June 1991.

Atkinson, Anthony A., Waterhouse, John H., Wells, Robert B., "A Stakeholder Approach to Strategic Performance Measurement", In ***Sloan Management Review***, Spring 1997, pp. 25-37

Ausindustries, Key Performance Indicators Manual: A Practical Guide for the Best Practice Development, Implementation and Use of KPIs, Business and Professional Publishing, May 1999

Austin, Robert D., Measuring and Managing Performance in Organizations, Dorset House Publishing, 1996

Baiman, Stanley, Rajan, Madhav V., "On the Design of Unconditional Monitoring Systems in Agencies", In ***The Accounting Review***, v.69, n.1, January 1994, pp. 217-229

Bagozzi, Richard P., Yi, Youjae, Phillips, Lynn W., "Assessing Construct Validity in Organizational Research", In ***Administrative Science Quarterly***, v.36, n.3, Sept 1991, pp. 421-458

Baker, Kathryn A., and Branch, Kristi M., "Concepts Underlying Organizational Effectiveness: Trends in the Organization and Management Science Literature", In ***Management Benchmark Study***, Office of Science, US Department of Energy, 2002

Baker, N.R., and Freeland, J.R., "Recent advances in R&D benefits measurement and project selection methods", In ***Management Science***, vol. 21, #3, 1975, pp.1164-75

Balkin, David B., Gomez-Mejia, Luis R., "Matching Compensation and Organizational Strategies", In ***Strategic Management Journal***, v.11, n.2, February 1990, pp. 153-169

Batty, J., (1988), 2nd Ed., Accounting for Research and Development. Gower Publishing Company Ltd., Aldershot, 238p.

Bauerschmidt, Alan, "Speaking of Strategy", In ***Strategic Management Journal***, v.17, n.8, Oct 1996, pp. 665-667

Bean, A.S., "Why some R&D Organizations are more productive than others?", In ***Research-Technology Management***, Jan-Feb. 1995

Beattie, C. J., and Reader, R. D., Quantitative Management in R&D, Chapman and Hall Ltd., 1971

Becker, Brian E., Huselid, Mark A., Ulrich, Dave, The HR Scorecard: Linking People, Strategy, and Performance, Harvard Business School Press, 2001

Berry, M.M.J., and Taggart, J.H., "Managing Technology and Innovation: A Review", In ***R&D Management***, Vol 24, #4, 1994, pp.341-353

Bett, Mike, "Quality with People or People with Quality?", In ***Engineering Management Journal***, August 1993, pp. 155-163

Birnbuam, Robert, "The Life Cycle of Academic Management Fads", In ***The Journal of Higher Education***, v.71, n.1, Jan-February 2000, p. 1-16

Blake, Stewart, P., Managing for Responsive Research and Development, W.H. Freeman and Company, 1978

Blenkhord, David L., Gaber, Brian, "The Use of 'Warm Fuzzies' to Assess Organizational Effectiveness", In ***Journal of General Management***, v.21, n.2, Winter 1995, pp. 40-51

Boer, F. P., The Valuation of Technology: Business and Financial Issues in R&D, NewYork: John Wiley & Sons, 1999

Bohan, George, "Whatever Happened to TQM? Or How a Good Strategy Got a Bad Reputation", In ***National Productivity Review***, Autumn 1998, pp. 13-16

Bolte, Keith A., "Productivity in the Engineering Disciplines", In ***National Productivity Review***, Spring 1986, pp. 134-141

Bozeman, B., and Melkers, J., Evaluating R&D Impacts, Methods and Practice, Boston: Kluwer Academic Publishers, 1993

Bozeman, Barry, Dietz, James S., "Strategic Research Partnerships: Constructing Policy Relevant Indicators", In ***Journal of Technology Transfer***, v.26, 2001, pp. 385-393

Bozeman, Barry, All Organizations are Public: Bridging Public and Private Organizational Theories, Jossey-Bass Publishers, 1987

Bradbury, F.R., Gallager, W.M., Suckling, C.W., "Qualitative Aspects of the Evaluation and Control of R&D Projects", In ***R&D Management***, vol. 3, #2, 1973, pp.49-57

Braver, Neal A.C., "Overcoming Resistance to TQM", In ***Research-Technology Management***, Sept.-October 1995, pp. 40-44

Brenner, Merrill S., Tao, John C., "You Can Measure External Research Programs", In ***Research-Technology Management***, May-June 2001, pp. 14-17

Brewer, Dominic J., Gates, Susan M., and Goldman, Charles A., In Pursuit of Prestige: Strategy and Competition in U.S. Higher Education, Transaction Publishers, 2002

Brook, Robert H., McGlynn, Elizabeth A., Shekkelle, Paul G., "Defining and Measuring Quality of Care: A Perspective from U.S. Researchers", In ***International Journal of Quality in Healthcare***, v.12, n.4, 2000, pp. 281-295

Brown, M. G., Svenson, R. A., "Measuring R&D Productivity", In ***Research-Technology Management,*** Vol. 41, #6, 1998, p. 30-36

Brown, Mark G., Keeping Score: Using the Right Metrics to Derive World Class Performance., Quality Resources, 1996.

Buderi, Robert, The Engines of Tomorrow: How the World's Best Companies are using Their Research Labs to Win the Future, Touchstone, 2000

Bugler, Stanley L., "Performance Management: The Foundation for High Performance Organization", In ***National Productivity Review***, Winter 1995/96, pp. 101-109

Bunch, Paul R., Schacht, Aaron L., "Modeling Resource Requirements for Pharmaceutical R&D", In ***Research-Technology Management***, January-February 2002, pp. 49-56

Byrne, George, "Ensuring Optimal Success with Six Sigma Implementations", In ***Journal of Organizational Excellence***, Spring 2003, pp. 43-50

Chaffee, Ellen E., "Three Models of Strategy", In ***The Academy of Management Review***, v.10, n.1, January 1985, pp. 89-98

Chatterji, Deb, and Davidson, Jeffery M., "Examining TQM's Legacies for R&D", In ***Research-Technology Management***, January-February 2001, pp. 10-21

Chein, Chen-Fu, "A Portfolio-Evaluation Framework for Selection R&D Projects", In ***R&D Management***, v.32, n.4, 2002, pp. 359-368

Chen, Chao C., Ford, Cameron M., Farris, George F., "Do Rewards Benefit the Organization? The Effects of Reward Types and the Perceptions of Diverse R&D Personnel", In ***IEEE Transactions on Engineering Management***, v.46, n.1, February 1999, pp. 47-55

Chew, Bruce W., "No-Nonsense Guide to Measuring Productivity", In ***Harvard Business Review***, Boston: Harvard Business School Press, January-February 1988, pp. 110-118

Chisea V., and Masella, C.,"Searching for an effective measure of R&D Performance", In ***Management Decision***, Vol. 34, #7, 1996, pp.49-57

Chotkowski, Marcel, and Follette, La, Quality in Science, The MIT Press, Cambridge, Massachusetts, 1982

Churchman, C. West, The Design of Inquiring Systems, New York: Basic Books, 1971

Clarke, Thomas E., Review of the R&D Management Literature Concerned with Career Management of Scientific Personnel, National Research Council of Canada, February 1996

Clarke, Thomas E., Unique Features of an R&D Work Environment and Research Scientists and Engineers, National Research Council of Canada, July 1996

Clarke, Thomas E., Focus and Expected Impact of Rewards, Recognition, and Incentives in Research Engineers and Scientists In Canadian Government Laboratories, National Research Council of Canada, December 1996

Coccia, Mario, "A Basic Model for Evaluating R&D Performance: Theory and Application in Italy", In ***R&D Management***, v.31, n.4, 2001, pp. 431-464

Cochran, M.A., "Investment Model for R&D Project Evaluation and Selection", In ***IEEE Transactions on Engineering Management***, vol. 18, #3, 1971, pp.89-100

Collier, Donald W., "Measuring the Performance of R&D Departments", In ***Research Management***, March 1977, pp. 30-34

Coombs, R., McMeekin, A., and Pybas, R., "Toward the development of benchmarking tools for R&D Project Management:, In ***R&D Management***, Vol. 28, #3, July 175-186, 1998

Cooper, R. G., Edgett, S. J., and Kleinschmidt, E. J., "Best practices for managing R&D Portfolios", In ***Research-Technology Management***., Vol. 41, #4, July-Aug, 20-33, 1998

Cordero, Rene, DiTomaso, Nancy, and Farris, George F., "Career Development Opportunities and Likelihood of Turnover among R&D Professionals", In ***IEEE Transactions on Engineering Management***, v.41, n.3, August 1994, pp. 223-233

Cotgrove, Stephen, "The Sociology of Science and Technology", In ***The British Journal of Sociology***, v.21, n.1, March 1970, pp. 1-15

COSEPUP, Evaluating Federal Research Programs: Research and Government Performance and Results Act (1993), Committee on Science, Engineering and Public Policy (COSEPUP), National Academies Press (NAP), 1999.

Crow, Michael, and Bozeman, Barry, Limited by Design, Columbia University Press, 1998

Cozzens, Susan E. "Quality of Life Returns from Basic Research." In ***The Contributions of Research to Society and Economy***. Washington: Brookings/ American Enterprise Institute, 1994.

Cozzens, Susan E. "Assessment of Fundamental Science Programs in the Context of the Government Performance and Results Act (GPRA)." MR-707.0-OSTP, Santa Monica: RAND, October 1995.

Daniels, Robin C., "Performance Measurement at the Sharp End", In ***Engineering Management Journal***, October 1995, pp. 211-214

Davidson, Jeffery M., Pruden, Ann L., "Quality Deployment in R&D Organizations", In ***Research-Technology Management***, January-February, 1996, pp. 49-55

Davis, J., Fusfeld, A., Scriven E., and Thritle, G., "Determining a Project's Probability of Success". In ***Research-Technology Management***, May-June, 2001, pp. 51-57

Debackere, Koenraad, Van Looy, Bart, Vliegen, Jan, "A Process View on Managing Quality During the Creation of Technical Innovations: Lessons from Field Research", In ***R&D Management***, v.27, n.3, 1997, pp. 197-221

Debra M. Amidon Rogers, "The Challenge of Fifth Generation R&D", In ***Research-Technology Management***, (July-August 1996).

Delaney, John T., Huselid, Mark A., "The Impact of Human Resources Management Practices on Perceptions of Organizational Performance", In ***The Academy of Management Journal***, v.39, n.4, August 1996, pp. 949-969

Demski, Joel S., "Implementation Effects of Alternative Performance Measurement Models in a Multivariate Context", In ***The Accounting Review***, v.46, n.2, April 1971, pp. 268-278

Diamond, Art, "What is Citation Worth?", In ***The Journal of Human Resources***, v. 21, no. 2, Spring 1986, pp. 200-215

Diamond, Art, "The Life-Cycle Research Productivity of Mathematicians and Scientists", In ***Journal of Gerontology***, v. 41, no. 4, 1986, pp. 520-525

Diamond, Art, "Science as a Rational Enterprise", In ***Theory and Decision***, v. 24, 1988, pp. 147-167

Diamond, Art, "Citation Counts for Nobel Prize Winners in Economics", In ***H.E.S. Bulletin***, v. 10, no. 1 Spring, 1988, pp. 67-70

Diamond, Art, "Economic Explanations of the Behavior of Universities and Scholars", In ***Journal of Economics Studies***, v. 20, no. 4/5. 1993, pp. 107-133

Diamond, Art, "The Economics of Science", In Knowledge and Policy: The ***International Journal of Knowledge Transfer and Utilization,*** v. 9, no. 2/3, Summar/Fall 1996, pp. 6-49

Diamond, Art, "The Career Consequences for a Scientist of a Mistaken Research Project", paper presented at ***Conference on "Evolution of Scientific Knowledge"*** Santa Fe Institute, 1998

Diamond, Robert M., Aligning Faculty Rewards With Institutional Mission: Statements, Policies, and Guidelines, Anker Publishing Company Inc., 1999

Diamond, Art, Constraints and Institutional Innovation in Science, Working Paper, University of Nebraska at Omaha, August 1999,

Diamond, Art, "Scientists Salaries and the Implicit Contracts Theory of Labor Markets", In ***International Journal of Technology Management***, v.22, n.7/8, 2001, pp. 688-697

Diamond, Art, and Toth, Robert J., The Determinants of Elections to the Presidency f American Economic Association, Working Paper, July 2002

Diamond, Art, "Edwin Mansfield's Contributions to the Economics of Technology", In ***Research Policy*** , v.32, no. 9, Oct. 2003, pp. 1607-1617.

Diamond, Art, "Zvi Griliches's Contributions to the Economics of Technology and Growth." In ***Economics of Innovation and New Technology*** , v.13, no. 4, June 2004, pp. 365-397.

Diamond, Art, "Zvi Ghriliches' Contributions to the Economics of Technology and Growth", In ***Economics, Innovation and New Technology***, v. 13, no. 4, June 2004, pp. 365-397

Diaz, Maria D. S., Gomez-Mejia, Luis R., "The Effectiveness of Organization-wide Compensation Strategies in Technology Intensive Firms", In ***The Journal of High Technology Management Research***, v.8, n.2, 1997, pp. 301-315

Dickson, David, "The virtue of 'science by numbers'" Editorial, Science and Development Network, August 2003, available at: http://www.scidev.net

Dill, David D., "Theory and Practice in the Staffing of R&D Laboratories", In ***R&D Management***, v.15, n.3, 1985, pp. 227-241

Domsch, M., Gerpott, T., and Jochum, E., "Peer Assessment in Industrial R&D Departments", In ***R&D Management***, v.13, n.3, 1983, pp. 143-153

Donabedian, Avedis, Explorations in Quality Assessment and Monitoring Vol-1, Health Administration Press, 1980

Drummond, Helga, "Beyond Quality", In ***Journal of General Management***, v.20, n.4, Summer 1995, pp. 68-77

Ehrbar, A.., EVA: The Real Key to Creating Wealth. New York: John Wiley & Sons, 1998

Ellis, Lynn W., Evaluation of R&D Processes: Effectivenes through Measurements. Boston, MA: Artech House, 1997

Ellis, Lynn W., Honig-Haftel, Sandra, "Reward Strategies for R&D", In ***Research-Technology Management***, March-April 1992, pp. 16-20

Endres, Al, Improving R&D Peformance the Juran Way, John Wiley and Sons, 1997

ERIMA, Evaluation of R&D Output, European Industrial Research Management Institute, Paris, Report # 47, 1985

ERIMA, Evaluation of R&D Projects. European Industrial Research Management Association, Paris, 1995

ERIMA, Centralized vs. Decentralized R&D, European Industrial Research Management Association Working Group 56, Paris, 1998

Ernst, Holger, Leptien, Christopher, and Vitt, Jan, "Inventors are not Alike: The Distribution of Patenting Output Among Industrial R&D Personnel", In ***IEEE Transactions of Engineering Management***, v.47, n.2, May 2000, pp. 184-199

Faulkner, T.W., "Applying Options thinking to R&D Valuation", In ***Research-Technology Management***, Vol. 39, # 3, May-Jun, 1996, pp. 50-56

Feeney, Anne-Marie, and Zairi, Mohammed, "TQM in Healthcare", In ***Journal of General Management***, v.22, n.1, Autumn 1996, pp. 35-43

Feller, Irwin, "The Who, What,…and How of Evaluating Science and Technology Programs, Presented at ***US/European Workshop on S&T Policy Evaluation***. Sponsored by School of Public Policy at Georgia Institute of Technology and the Fraunhofer Institute for Systems and Innovation Research (ISI), Germany, Sept. 2000

Flamholtz, Eric G., "Towards a Psycho-Technical Systems Paradigm of Organizational Measurement", In ***Decision Sciences***, v.10, 1979, pp. 71-84

Foster, Ted N., "Making R&D More Effective at Westinghouse", In ***Research-Technology Management***, January-February 1996, pp. 31-37

Foster, R. N., "Assessing Technology Threats", In ***Research Management***, Vol. 29, #4, 1986, pp.17-20

Foster, R.N., Linden, L.H., Whitney, R.L., and Kantrow, A.M., "Improving the return on R&D-I", In ***Research-Technology Management***, Jan-Feb., 1985

Foster, R.N., Linden, L.H., Whitney, R.L., and Kantrow, A.M., "Improving the return on R&D-II", In ***Research-Technology Management***, March-April, 1985

Foster, R. N., "Boosting the payoff from R&D", In ***Research Management***, Vol.25, #1, 1982, pp.22-27

Francis, P.H., "Putting Quality back into the R&D Process", In ***Research-Technology Management***, July-August 1992

Freeman, Richard B., Kleiner, Morris M., "Who Benefits Most from Employee Involvement: Firms or Workers?", In ***The American Economic Review***, v.90, n.2, May 2000, pp. 219-223

Friedlob, George T., Schliefer, Lydia, L. F., and Plewa Jr., Franklin J., Essentials of Corporate Performance Measurement, John Wiley and Sons Inc., 2002

Gaither, Gerald H. (Ed.), Assessing Performance in the Age of Accountability, Jossey-Bass Inc., 1995

Galbraith, Craig, Schendel, Dan, "An Empirical Analysis of Strategy Types", In ***Strategic Management Journal***, v.4, n.2, April-June 1983, pp. 153-173

Gandhi, T. K., Francis, E. C., Puopolo, A. L., Burstin, H. R., Haas, J. S., Brennan, T. A., "Assessing the Comparability of Various Measures of the Quality of Ambulatory Care", In ***Medical Care***, v.40, n.2, 2002, pp. 155-165

GAO, Measuring Performance: Strengths and Limitations of Research Indicators, Government Accounting Office, GAO/RCED-97-91, March 1997

GAO, Measuring Performance: Challenges in Evaluating Research and Development, Government Accounting Office, Congressional Testimony, GAO/RCED-97-130, April 1997

Garside, Pam, "Organizational Context for Quality: Lessons from the Fields of Organizational Development and Change Management", In ***Quality in Healthcare***, v.7, December 1998, pp. S8-S15

Gauch, Hugh G. Jr., Scientific Method in Practice, Cambridge University Press, 2003

Geisler, Eliezer, The Metrics for Science and Technology, Quorum Books, 2000

Geles, Claude; Lindecker, Gillesl; Month, Mel; and Roche, Christian, Managing Science: Management for R&D Laboratories, John Wiley and Sons, 2000

Gibbs, Michael, Returns to Skill & Personnel Management: U.S. DoD Scientists and Engineers, Working Paper, Graduate School of Business, University of Chicago, Sept. 2001

Gibbs, Michael, and Levenson, Alec, The Economic Approach to Personnel Management, Draft Paper, Sept. 2000

Ginsberg, Ari, "Measuring and Modeling Changes in Strategy: Theoretical Foundations and Empirical Directions", In ***Strategic Management Journal***, v.9, n.6, November-December 1988, pp. 559-575

Ginsberg, Ari, "Operationalizing Organizational Strategy: Towards an Integrative Framework", In ***The Academy of Management Review***, v.9, n.3, July 1984, pp. 548-557

Gleason, John M., Barnum, Darold T., "Toward Valid Measures of Public Sector Productivity: Performance Measures in Urban Transit", In ***Management Science***, v.28, n.4, April 1982, pp. 379-386

Godin, B., Defining R&D: Is Research Always Systematic? , Working Paper No. 7, Project on History and Sociology of S&T Statistics, 2000

Godin, B., Tradition and Innovation: The Historical Contingency of R&D Statistical Classifications, Working Paper No. 11, Project on History and Sociology of S&T Statistics, 2001

Godin, B., A Note on the Survey as an Instrument for Measuring Science and Technology, Working Paper No. 18, Project on History and Sociology of S&T Statistics, 2002

Godin, B., "The Numbers Makers: Fifty Years of Science and Technology Official Statistics", In ***Minerva***, v.40, 2002, pp. 375-397

Godin, B., The Who, What, Why and How of S&T Measurement, Working Paper, INRS, Montreal, undated

Goldman, Charles A., Massay, William F., The PhD Factory: Training and Employment of Science and Engineering Doctorates n the United States, Anker Publishing Company Inc., 2001

Gooderham, Gay, "The Top 10 Lessons for Implementing Performance Management Systems", In ***Journal of Cost Management***, January-February 2001, pp. 29-33

Gormley Jr., William T., and Weimer, David L., Organizational Report Cards, Harvard University Press, Cambridge: Massachusetts, 1999

Grady D., Fincham T., "Making R&D Pay", In ***Research-Technology Management,*** March-April 1991

Grant, Robert, "AMR Captures TQM – Essence Escapes", In ***The Academy of Management Review***, v.20, n.1, January 1995, pp. 11-15

Gratton, Lynda, "How can we Predict Management Potential in Research Scientists?, In ***R&D Management***, v.17, n.2, 1987, pp. 87-97

Graves, S.B., Ringuest, J.L., "Evaluating Competing R&D Investments", In ***Research-Technology Management***, vol. 34, #4, 1991, pp. 32-35

Griffith, A. and Page, A.L., "PDMA Success Measurement Project: Recommended Measures for Product Development Success and Failure", In ***Journal of Product Innovation Measurement***, Vo. 13, 1996, pp.478-496

Griliches, Zvi, Hybrid Corn: "An Exploration in the Economics of Technological Change", In ***Econometrica***, October 1957, Vol. 25(4), 501-522

Griliches, Zvi, "Hybrid Corn and the Economics of Innovation", In ***Science***, l32, 275-280, July 29, 1960.

Gronow, Luukonen T., "Scientific Research Evaluation: A Review of Methods and Various Contexts of their Application", In ***R&D Management***, v. 17, n.3, 1987, pp. 207-221

Gupta, Praveen, Six Sigma Business Scorecard: Ensuring Performance for Profit, McGraw-Hill, 2004

Gupta, Ashok K., Wilemon, David, Atuahene-Gima, Kwaku, "Excelling in R&D", In Research-Technology Management, May-June 2000, pp. 52-58

Gustin, Bernard H., "Charisma, Recognition, and the Motivation of Scientists", In ***American Journal of Sociology***, v.78, n.5, March 1973, pp. 1119-1134

Hackman, Richard J., Wagerman, Ruth, "Total Quality Management: Empirical, Conceptual, and Practical Issues", In ***Administrative Science Quarterly***, v.40, n.2, July 1995, pp. 309-342

Hahn, Gerard J., Hill, William J., Hoerl, Roger W., Zinkgraf, Stephen A., "The Impact of Six Sigma Improvement—A Glimpse into the Future of Statistics", In ***The American Statistician***, v.53, n.3, August 1999, pp. 208-215

Halachmi, Arie, and Bouckaert, Geert, Organizational Performance and Measurement in the Public Sector: Toward Service, Effort, and Accomplishment Reporting, Quorum Books, 1997

Halachmi, Arie, "Business Process Reengineering in the Public Sector: Another Frog to Fly?", In ***National Productivity Review***, Summer 1996, pp. 9-29

Halachmi, Arie, "Mandated Performance Measurement: A Help or a Hindrance?", In ***National Productivity Review***, Spring 1999, pp. 59-67

Halperin, Michael R., Chakrabarti, Alok K., "Firm and Industry Characteristics Influencing Publications of Scientists in Large American Companies", In ***R&D Management***, v.17, n.3, 1987, pp. 167-173

Hart, Stuart L., "An Integrative Framework for Strategy-Making Processes", In ***The Academy of Management Review***, v.17, n.2, April 1992, pp. 327-351

Hart, Stuart, Banbury, Catherine, "How Strategy-Making Processes Can Make a Difference", In ***Strategic Management Journal***, v.15, n.4, May 1994, pp. 251-269

Hatten, Kenneth J., Rosenthal, Stephen R., "Why-and-How-To Systematize Performance Measurement", In ***Journal of Organizational Excellence***, Autumn 2001, pp. 59-73

Hauser, Scott J., and McEnaney, Celia (Eds.), Strategies for Competitiveness in Academic Research, Committee on Science, Engineering, and Public Policy, AAAS, 2000

Hauser, John R., Metrics Thermostat, MIT Sloan School Working Paper, December 2000

Hauser, John R., and Zettelmeyer, Florian, "Metrics to Evaluate R, D & E", In ***Research-Technology Management***, July-August 1997, pp. 32-38

Hauser, J.R., Clausing, D., "The House of Quality", In ***Harvard Business Review***, Vol.66, #3, 1988, pp.63-73

Hedge, D. G., "Quality in R&D -- Culture and Practice", In ***Total Quality Management, Proceedings of the 2nd International Conference,*** IFS Publications, 1990

Hendricks, Kevin, Menor, Larry, and Wiedman, Christine, Adopting of the Balanced Scorecard: A Contingency Variable Analysis, Working Paper, Richard Ivey School of Business, University of Western Ontario, September 2004

Hollins, Carol S. (Ed), Containing Costs and Improving Productivity in Higher Education, Jossey-Bass Publishers, 1992

Holloway, Jacky, Lewis, Jenny, and Mallory, Geoff (Eds.), Performance Measurement and Evaluation, Sage Publications, 1995

Howard, David, "An Approach to Improving Management Performance—Part 1", In ***Engineering Management Journal***, April 1992, pp. 63-69 [Howard, 1992a]

Howard, David, "An Approach to Improving Management Performance—Part 2", In ***Engineering Management Journal***, June 1992, pp. 123-130 [Howard, 1992b]

Hurley, John, Organization and Scientific Discovery, Jossey-Bass Inc., 1997

Hurley, John, Scientific Research Effectiveness, Kluwer Academic Publishers, 2003

Hodge Jr., Melville H., "Rate Your Company's Research Productivity", In ***Harvard Business Review***, Boston: Massachusetts, January 26, 1959, pp. 109-122

Igbaria, Magid, Kassicieh, Suleiman K., and Silver, Martin, "Career Orientations and Career Success among Research, and Development, and Engineering Professionals", In ***Journal of Technology Management***, v. 16, 1999, pp. 29-54

Inkpen, Andrew, Choudhary, Nandhan, "The Seeking of Strategy Where It is Not: Towards a Theory of Strategy Absence", In ***Strategic Management Journal***, v.16, n.4, May 1995, pp. 313-323

IRI, "Industrial Research Institute's R&D Trends Forecast for 2002", In ***Research-Technology Management***, January February 2002, pp. 16-20

Iselin, Errol R, Mia, Lokman, Sands, John, The Effects of Balanced Scorecard and Related Systems on Organizational Performance, Working Paper, Griffith University, undated

Ittner, Christopher D., Larcker, David F., Rajan, Madhav V., "The Choice of Performance Measures in Annual Bonus Contracts", In ***The Accounting Review***, v.72, n.2, April 1997, pp. 231-255

Jamsen, Miikka, Suomala, Petri, and Paranko, Jari, What is being Measured in R&D: Contradictions between the need and the Practice, Working Paper, Tempere University of Technology, Finland, 2001

Johnson, Albert, "Six Sigma in R&D", In ***Research-Technology Management***, March-April 2002, pp. 12-16

Johnson, Albert, and Swisher, Beth, "How Six Sigma Improves R&D", In ***Research-Technology Management***, March-April, 2003, pp. 12-15

Jordan, Gretchen B., Malone, Elizabeth L., "Performance Assessment", In ***Management Benchmark Study***, Office of Science, US Department of Energy, 2002

Jordan, Gretchen B., Streit, Devon, and Binkley, Stephen J., "Assessing and Improving the Effectiveness of National Research Laboratories", In ***IEEE Transactions on Engineering Management***, v.50, n.2, May 2003, pp. 228-235

Kaplan, Robert S., and Norton, David P., "The Balanced Scorecard: Measures that Drive Performance." In ***Harvard Business Review***, January-February 1992, pp. 71-79.

Kaplan, Robert S., and David P. Norton. "Putting the Balanced Scorecard to Work" In ***Harvard Business Review***, Sept.-October 1993, pp. 134-147

Kaplan, Robert S., and David P. Norton. "Using the Balanced Scorecard as a Strategic Management System." In ***Harvard Business Review***, January-February 1996, pp. 75-85.

Kaplan, Robert S., Norton, David P., The Balanced Scorecard, Harvard Business School Press, 1996

Kaplan, Robert S., and Norton, David P., "Having Trouble with Your Strategy? Then Map It." In ***Harvard Business Review***, September-October 2000, pp. 3-11

Kaplan, Robert S., Norton, David P., The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment, Harvard Business School Press, 2001

Kaplan, Robert S., Norton, David P., The Balanced Scorecard Executive Summit 2002, Cambridge, Massachusetts, December 2002

Kaplan, Robert S., Norton, David P., Strategy Maps: Converting Intangible Assets into Tangible Outcomes, Harvard Business School Press, 2004

Katz, Ralph (Ed.), Managing Professionals in Innovative Organizations, Ballinger Publishing Company, 1988

Katz, Sylvan J., "Scale-independent Indicators and Research Evaluation", In ***Science and Public Policy***, February 2000, pp. 23-36

Kearney, Richard C., and Berman, Evan M. (Eds.), Public Sector Performance: Management, Motivation, and Measurement, Westview Press, 1999

Keehley, P., Medlin, S., MacBride, S., Longmire, L., Benchmarking for Best Practices in the Public Sector: Achieving Performance Breakthroughs in Federal, State, and Local Agencies, Jossey-Bass Inc., 1997

Keiser, Bruce A., Blake, Natalie, R., "How Nalco Revitalized Its Quality Process for R&D", In ***Research-Technology Management***, May-June 1996, pp. 23-29

Keller, Robert T., Julian, Scott D., and Kedia, Ben L., "A Multinational Study of Work Climate, Job Satisfaction, and Productivity of R&D Teams", In ***IEEE Transactions on Engineering Management***, v.42, n.1, February 1996, pp. 48-55

Kersselen-van Drongelen, I.C., and Cook, Andrew, "Design Principles for the Development of Measurement Systems for Research and Development Processes", In ***R&D Management***, v.27, n.4, 1997, pp. 345-357

Kersselen-van Drongelen, I.C., Systematic Design of R&D Performance Measurement Systems. PhD Thesis, University of Twente, Enschede, The Netherlands, 1999

Kerssens-van Drongelen, I.C., Nixon, W. A., Pearson, A. W., "Performance Measurement in Industrial R&D", In ***International Journal of Management Reviews***, v.2, n.2, 2000, pp. 111-144

Kim, Bowon, and Oh Heungshik, "Economic Compensation Compositions Preferred by R&D Personnel of Different R&D Types and Intrinsic Values", In ***R&D Management***, v.32, n.1, 2002, pp. 47-59

Kleinknecht, Alfred, and Donald Bain, Eds. New Concepts in Innovation Output Measurement. Houndsmill: Palgrave Macmillan, 1993.

Kostoff, Donald, N. The Handbook of Research Impact Assessment, Unpublished manuscript, Office of Naval Research, 1997.

Kostoff, Ronald N. Research Program Peer Review: Principles, Practices, Protocols. Arlington, VA, 1999

Kostoff, Donald N. Science and Technology Metrics. Unpublished manuscript, Office of Naval Research, 2000.

Krogh, Lester C., Prager, Julliane H., Soreson, David P., Tomlinson, John D., "How 3M Evaluates its R&D Programs", In ***Research-Technology Management***, November-December 1988, pp. 10-14

Krumholz, H. M., Rathore, S. S., Chen, J., Wang, Y., and Radford, M. J., "Evaluation of a Consumer-Oriented Internet Health Care Report Card", In ***Journal of American Medical Association***, v. 287, n.10, pp. 1277-1287

Kuwahara, Y., Takeda, Y., "A Managerial Approach to Research and Development Cost Effectiveness Evaluation", In ***IEEE Transactions on Engineering Management***, Vol. 37, n.2, 1990

Kuhn, Thomas, The Structure of Scientific Revolutions, The University of Chicago Press, 1962

Lammers, J. C., Cretin, S., Gilman, S., and Calingo, E., "Total Quality Management in Hospitals", In ***Medical Care***, v. 34, n.5, 1996, pp. 463-478

Landy, Frank, Zedeck, Sheldon, Cleveland, Jeanette (Eds.), Performance Measurement and Theory, Lawrence Erlbaum Associates Publishers, 1983

Larsson, Rikard, "Case Survey Methodology: Qualitative Analysis of Patterns Across Case Studies", In ***The Academy of Management Journal***, v.36, n.6, December 1993, pp. 1515-1546

Latham, Gary P., Steele, Timothy P., "The Motivational Effects of Participation versus Goal Setting on Performance", In ***The Academy of Management Journal***, v.26, n.3, Sept 1983, pp. 406-417

Latour, B., Woolgar, S., Laboratory Life: The Construction of Scientific Facts, Princeton University Press, 1986

Lawler, Edward A., Motivation in Work Organizations, Jossey-Bass Inc., 1994

Lawler, Edward E., Rewarding Excellence: Pay Strategies for the New Economy, Jossey-Bass Inc., 2000

Lazear, Edward P., "Performance Pay and Productivity", In ***The American Economic Review***, December 2000, pp. 1346-1361

LeBonte, Thomas J., Building a New Performance Vision, The American Society for Training and Development, 2001

Lee, Dennis M. S., "Job Challenge, Work Effort, and Job Performance of Young Engineers: A Causal Analysis", In ***IEEE Transactions on Engineering Management***, v.39, n.3, August 1992, pp. 214-226

Leifer, Robert, Burke, William J., "Organizational Activity Analysis: A Methodology for Analyzing and Improving Technical Organizations", In ***IEEE Transactions on Engineering Management***, v.41, n.3, August 1994, pp. 234-244

Loch, Christoph A., Tapper, Staffan U.A., "Implementing a Strategy-Driven Performance Measurement System for an Applied Research Group", In ***The Journal of Product Innovation Management***, v. 19, 2002, pp. 185-198

Loiuse, Searshore K., Blumenthal, David, Gluck, Michael E., and Stoto, Michael A., "Entrepreneurs in Academe: An Exploration of Behaviors among Life Scientists", In ***Administrative Science Quarterly***, v.34, n.1, March 1989, pp. 110-131

Luzi, Andrew D., Mackenzie, Kenneth D., "An Empirical Study of Performance Information Systems", In ***Management Science***, v.28, n.3, March 1982, pp. 243-259

Magnusen, Karl O., "Faculty Evaluation, Performance, and Pay: Application and Issues", In ***The Journal of Higher Education***, v.58, n.5, Sept-Oct. 1987, pp. 516-529

Maisel, L. S., "Performance Measurement: The Balanced Scorecard Approach", In ***Journal of Cost Management***, Summer 1992, pp. 47-52

Malina, Mary, and Selto, Frank, "Communicating and Controlling Strategy: An Empirical Study of Effectiveness of the Balanced Scorecard", In Journal of Management AccountingResearch, v.13, 2001

Maltz, Alan C., Shenkar, Aaron J., Reilly, Richard R., "Beyond the Balanced Scorecard: Refining the Search for Organizational Success Measures", In ***Long Range Planning***, v.36, 2003 pp. 187-204

Mangemati, V., "PhD Job Market: Professional Trajectories and Incentives During the PhD", In ***Research Policy***, v. 29, 2000, pp. 741-756

Mansfield, Edwin, "How Economists See R&D", In ***Harvard Business Review***, November-December, 1981, pp. 98-106

Marien, Bruce A., "Putting TQM Into the Performance Review Process", In ***Research-Technology Management***, Sept-October 1992, pp. 39-43

Martel, Leon, "The Principles of High Performance—And How to Apply them", In ***Journal of Organizational Excellence***, Autumn 2002, pp. 49-59

Mason, Richard O., "Basic Concepts for Designing Management Information Systems", In Mason, Richard O., Swanson, E. Burton, ***Measurement for Management Decision***, Addison-Wesley Publishing Company, 1981

Mason, Richard O., Swanson, E. Burton, Measurement for Management Decision, Addison-Wesley Publishing Company, 1981

May, Colin, and Pearson, Alan M., "Total Quality in R&D", In ***Journal of General Management***, v.18, n.3, Spring 1993, pp. 1-22

McClenahen, John S., "The Jack Welch of Washington", In ***Industry Week***, February 21, 2001

McGlynn, Elizabeth A., "Six Challenges In Measuring The Quality of Healthcare", In ***Health Affairs***, v.16, n.3, May-June 1997

McGlynn Elizabeth A., Brook, R. H., "Evaluating the Quality of Care", In R. M. H., Rice and G. F. Kominski (Eds.), ***Changing the U.S. Healthcare System: Key Issues in Health Services Policy and Management*** 2nd Ed., New York: Jossey-Bass Inc., 2001

McGrath, M. M., Romeri, M. N., "The R&D Effectiveness Index: A Metric for Product Development Performance", In ***Journal of Product Innovation Mgmt***., Vol. 11, #3, 1994, pp.213-220

Mclaughlin, Curtis P., Stratman, Jeffrey K., "Improving the Quality of Corporate Technical Planning: Dynamic Analogues of QFD", In ***R&D Management,*** Vol. 27 No. 3, 1997

Mclaughlin, Gregory C., Total Quality in Research and Development, St. Lucie Press, 1995

McMillan, G. Steven, Deeds, David L., "The Role of Reputation in the Recruitment of Scientists", In ***R&D Management,*** v.28, n.4, 1998, pp. 299-304

McKenzie, Richard B., Lee, Dwight R., Managing Through Incentives: How to Develop a More Collaborative, Productive, and Profitable Organization, Oxford University Press, 1998

McNair, C. J., Lynch, R. L., and Cross K. F., "Do Financial and Non-financial Performance Measures Have to Agree?", In ***Management Accounting***, November 1990, pp. 28-35

McNeil, B. J., "Shattuck Lecture: Hidden Barriers to Improvement in the Quality of Care", In ***The New England Journal of Medicine***, v.345, n.22, November 2001, pp. 1612-1620

Meinhart, Wayne A., Pederson, John A., "Measuring the Performance of R&D Professionals", In ***Research-Technology Management***, July-August 1989, pp. 19-21

Metz, Edmund J., "Using Strategic Design Concepts to Create A Best-In-Class Performance Management System", In ***National Productivity Review***, Spring 2000, pp. 51-62

Meyer, C., "How the right measures help teams to excel", In ***Harvard Business Review***, Boston: Massachusetts, May-June 1994

Miller, C. Chet, Cardinal, Laura B., "Strategic Planning and Firm Performance: A Synthesis of More than Two Decades of Research", In ***The Academy of Management Journal***, v.37, n.6, December 1996, pp. 1649-1665

Miller, J. A., Norling, P. M., Collette, J. W., "Research/Technology Management", In ***Kirk-Othmer Encyclopedia of Chemical Technology*** (Fourth Ed.), v.21, John Wiley and Sons, 1997

Miller, William L., Morris, Langdon, Fourth Generation R&D: Managing Knowledge, Technology, and Innovation, Wiley and Sons Inc., 1999

Miller, R., "Applying Quality Practices to R&D", In ***Research-Technology Management***, March-April, 1995, pp.47-53

Miller, Roger, "The Influence of Primary Task on R&D Laboratory Evaluation: A Comparative Bibliometric Analysis", In ***R&D Management***, v.22, n.1, 1992, pp. 3-19

Mintzberg, Henry, "Patterns in Strategy Formulation", In ***Management Science***, v.24, n.9, May 1978, pp. 934-948

Mitchell, Terence R., "An Evaluation of the Validity of Correlational Research Conducted in Organizations", In ***The Academy of Management Review***, v.10, n.2, April 1985, pp. 192-205

Mohrman, Allan M., Mohrman, Susan A., "Self-Designing a Performance Measurement System", In Adler, Nicholas, Shani, A.B., and Styhre, Alexander (Eds.), ***Collaborative Research in Organizations: Foundations for Learning, Change, and Theoretical Development***, Sage Publications, 2004

Moore, Gordon E., "Cramming More Components onto Integrated Circuits", In ***Electronics***, v.38, n.8, April 19, 1965

Morley, Elaine, Bryant, Scott P., and Hatry, Harry P., Comparative Performance Measurement, The Urban Institute Press, Washington D.C., 2001

Mosakowski, Elaine, "Strategy-Making Under Causal Ambiguity: Conceptual Issues and Empirical Evidence", In ***Organization Science***, v.8, n.4, July-August 1997, pp. 414-442

Moser, M.R., "Measuring Performance in R&D Settings", In ***Research-Technology Management***, vol. 28, n.5, 1985

Mulkay, M. J., Williams, Anthony T., "A Sociological Study of a Physics Department", In ***The British Journal of Sociology***, v. 22, n.1, March 1971, pp. 68-82

Murray, T. J., "Quality in R&D, Meeting the New Challenge of Quality", In ***Research Management***, 1987

Narayanan, V. K., Fahey, Liam, "The Micro-Politics of Strategy Formulation", In ***The Academy of Management Review***, v.7, n.1, January 1982, pp. 25-34

Nath, Deepika, and Sudharshan, D., "Measuring Strategic Coherence Through Pattern of Strategic Choices", In ***Strategic Management Journal***, v.15, n.1, January 1994, pp. 43-61

Neely, Andy, Marr, Bernard, Roos, Goran, Pike, Stephen, and Gupta, Oliver, "Towards the Third Generation of Performance Measurement", In ***Controlling***, April 2003

Neely, Andy, Adams, Chris, Kennerly, Mike, The Performance Prism: The Scorecard for Measuring and Managing Business Success, Financial Times – Prentice Hall, 2002

Nelson, Richard, "The Economics of Invention: A Survey of Literature", In ***Journal of Business***, v. 32., n.2, pp. 101-127, 1959

Nielsen, Steen, and Sorensen, Rene, "Motives, Diffusion, and Utilization of the Balanced Scorecard in Germany", In ***International Journal of Auditing and Performance Evaluation***, v.1, n.1, 2004

Nevin, Paul R., Balanced Scorecard: Step-by-Step for Government and Non-Profit Agencies, John Wiley and Sons Inc., 2003

Noll, Roger G., Challenges to Research Universities, Brookings Institution Press, 1998
Lawrence, Peter A., "Politics of Publication", In ***Nature***, 422, March 20, 2003, pp. 259 – 261

Norling, Parry M., High Performance Work Systems in Technical Organizations, A presentation, 1993

NRC, Assessment of the U.S. Army Natick Research, Development, and Engineering Center, National Research Council, National Academies Press, 1997

NRC, World Class Research and Development, National Research Council, National Academies Press, 1996

Oberg, James, "Houston, We Have a Problem", In ***New Scientist***, v.166, Issue.2234, 15 April 2000, p. 26

OECD (Organization of Economic Cooperation for Development), In Search of Results: Performance Management Practices, 1997

Olve, Nils-Goran, J. Roy, and M. Wetter. Performance Drivers: A Practical Guide to Using the Balanced Scorecard. West Sussex: John Wiley & Sons, 1999.

OMB, Instructions for Program Assessment Rating Tool, Office of Management of the Budget (OMB), Washington DC, 2004

Omta, S.W.F., Critical Success Factors in Biomedical Research and Pharmaceutical Innovation, Kluwer Academic Publishers, 1995

Osama, Athar, Held, Bruce, Menthe, Lance, Designed to Deliver: Using a Systematic Organizational Framework to Create Army's New RDE Organization from First Principles, DRR-3000-A, Santa Monica: RAND, March, 2003

Osama, Athar, "Mission-2-Metrics: A Balanced Scorecard-based R&D Measurement System to Enable US Army's RDE Reorganization", AB-643-A, Santa Monica: RAND, September 2002

Pace, Larry A., Kelly, Eileen P., "TQM at Xerox: Lessons Worth Duplicating", In ***International Journal of Technology Management***, v.16, n.4/5/6, 1998, pp. 326-335

Packer, Michael B., "Analyzing productivity in R&D Organizations", In ***Research-Technology Management,*** Jan-Feb. 1983

Pappas, R.A., Remer, D.S., "Measuring R&D Productivity", In ***Research-Technology Management***, May-June 1985

Parish, Truman, D., "Technology Value Pyramid", In ***Assessing the Value of Research in Chemical Sciences***, Commission on Physical Sciences, Mathematics, and Applications (CPSMA), National Academy Press, 1998, pp. 50-55

Parkinson, Steven, Rodney, McAdam, and Henderson, Joan, "Organizational Change Evaluation" Towards a More Subjective Approach", In ***Journal of General Management***, v.25, n.2, Winter 1999, pp. 59-79

Patino, Hugo, "Applying Total Quality to R&D at Coors Brewing Company", In ***Research-Technology Management***, Sept-October, 1997, pp. 32-36

Patterson, William C., "Evaluating R&D Performance at Alcoa Laboratories", In ***Research Management***, March-April 1983, pp. 23-27

Payne, R., "Individual Differences & Performance Amongst R&D Personnel: Some Implications for Management Development", In ***R&D Management***, v.17, n.3, 1987, pp. 153-165

Pearson, A. W., Nixon, W. A., Kerssens-van Drongelen, I. C., "R&D as a Business—What are the Implications for Performance Measurement", In ***R&D Management***, v.30, n.4, 2000, pp. 355-366

Pearson, Alan W., Vaughan, Nicci, and Butler, Jeff, "The Implementation of TQM in R&D", In ***International Journal of Technology Management***, v. 16, n.4-6, 1998, pp. 405-431

Pelikan, Jaroslav, The Idea of the University – A Reexamination, Yale University Press, 1992

Pelz, Donald C., Andrews, Frank M., Scientists in Organizations: Productive Climates for Research and Development, John Wiley and Sons, 1966

Pencikowski, Paul, "Faster, Better, Cheaper? A Critical View", 2000

Pinto, Jeffery K., Slevin, Dennis P., "Critical Success Factors in R&D Projects", In ***Research-Technology Management***, January-February, 1989, pp. 31-35

Popper, Steven W. Economic Approaches to Measuring the Performance and Benefits of Fundamental Science. MR-708.0, Santa Monica: RAND, October 1995.

Porter, M., "What is Strategy?", In ***Harvard Business Review***, November-December, 1996, p. 62

Porter, A.C., et al. (1980), A Guidebook for Technology Assessment and Impact Analysis, New York: North Holland.

Powell, Thomas C., "Organizational Alignment as Competitive Advantage", In ***Strategic Management Journal***, v.13, n.2, February 1992, pp. 119-134

Presley, Adrian and Liles, Donald, "R&D Validation Planning: A Methodology to Link Technical Validations to Benefits Measurement", In ***R&D Management***, Vol. 30. No. 1, 2000

Price, R.C., Gaskill, G. P., "Total Quality Management in Research – Philosophy and Practice", In ***Total Quality Management, Proceedings of the 2nd International Conference***, IFS Publications, 1990

Purdon, William A. B., "Increasing R&D Effectiveness: Researchers as Business People", ***In Research-Technology Management***, July-August 1996, pp. 48-56

Quinn, J.B., "How to Evaluate Research Output", In ***Harvard Business Review***, March-April 1960

Ranftl, Robert M., R&D Productivity—An Investigation of Ways to Enhance and Improve Productivity in Technology-based Organizations, Hughes Aircraft Company, Culver City, 1978

Ranftl, Robert M., R&D Productivity, In ***Chemtech***, November 1980, pp. 661-669

Ransley, Derek L., "Do's and Don'ts of R&D Benchmarking", In ***Research-Technology Management***, Sept.-October, 1994, pp. 50-56

Rampersad, Herbert K., Total Performance Scorecard: Redefining Management to Achieve Performance with Integrity, Butterworth-Heinemann, May 2003

Reger, Rhonda K., Gustafson, Loren T., Demarie, Samuel M., Mullane, John V., "Reframing the Organization: Why Implementing Total Quality is Easier Said than Done", In ***The Academy of Management Review***, v.19, n.3, July 1994, pp. 565-584

Reskin, Barbara F., "Scientific Productivity and the Reward Structure of Science", In ***American Sociological Review***, v.42, n.3, January 1977, pp. 491-504

Risher, Howard, "Compensating Today's Technical Professional", In ***Research-Technology Management,*** January-February 2000, pp. 50-56

Robb, W.L., "How Good is our Research", ***In Research-Technology Management***, March-April, 1991

Roberts E. B., "Benchmarking the Strategic Management of Technology-I", In ***Research-Technology Management***, January-February 1995, pp. 44-56

Roberts E. B., "Benchmarking the Strategic Management of Technology-II", In ***Research-Technology Management***, March-April 1995, pp. 18-26

Roberts, E.B., "Managing Invention and Innovation: What we've learned. ", In ***Research-Technology Management***, Vol. 31, Jan-Feb 1998

Roberts, E.B., Managerial Determinant of Industrial R&D Performance: An Analysis of Global Chemicals/Materials Industry, MIT Sloan School of Management Working Paper, 1994

Robinson, William W., "A Systems Approach to Management", In ***Engineering Management Journal***, August 1996, pp. 172-176

Robinson, William W., "There is Method In't", In ***Engineering Management Journal***, June 1999, pp. 100-104

Root-Bernstein, Robert S., "Who Discovers and Invents", In ***Research-Technology Management***, January-February 1989, pp. 43-50

Roth, William, and Capuano, Terry, "Systemic Versus Nonsystemic Approaches to Quality Improvement", In ***Journal of Organizational Excellence***, Spring 2001, pp. 57-64

Roussel, Philip A., Saad, Kamal N., Erickson, Tamara J. , Third Generation R&D: Managing the Link to Corporate Strategy, Harvard Business School Press, 1991

Rubin, H. R., Pronovost, P., Diette, G. B., "From a Process of Care to a Measure: The Development and Testing of a Quality Indicator", In ***International Journal of Quality of Healthcare***, v. 13, n.6, 2001, pp. 489-496

Rumsfeld, Donald, H., Annual Report to the President and Congress, Department of Defense, 2004

Russell, Jennifer, Evaluating Early Evidence on Implementation of Accountability Report Cards, Working Paper, WR-202-EDU, Santa Monica: RAND, November 2004

Ryan, George R., "Dealing with the Survivors of A Reorganization", In ***Research-Technology Management***, March-April 1989, pp. 40-42

Rycroft, Robert W., Kash, Don E., The Complexity Challenge: Technological Innovation for the 21st Century, Cassell Academic, 1999

Ryder, Frank H., "The Continuous Improvement of Organizational System Performance", In ***Engineering Management Journal***, August 1992, pp. 183-188

Ryan, George R., "Dealing with the Survivors of A Reorganization", In ***Research-Technology Management***, March-April 1989, pp. 40-42

Saleh, Shoukry D., Wang, Clement K., "The Management of Innovation: Strategy, Structure, and Organizational Climate", In ***IEEE Transactions on Engineering Management***, v.40, n.1, February 1993, pp. 14-21

Saxberg, Borje O., Slocum, John W., "The Management of Scientific Manpower", In ***Management Science***, v.14, n.8, April 1968, pp. B473-B489
Schumman, Paul A., Ransley, Derek L., Prestwood, Donna C. L., "Measuring R&D Performance", ***Research-Technology Management***, May-June 1995, pp. 45-54

Segev, Eli, "Strategy, Strategy-Making, and Performance – An Empirical Investigation", In ***Management Science***, v.33, n.2, February 1987, pp. 258-269

Shapira, Zur, Griffith, Terri L., "Comparing the Work Values of Engineers with Manageers, Production, and Clerical Workers: A Multi-variate Analysis", In ***Journal of Organizational Behavior***, v.11, n.4, July 1990, pp. 281-292

Sheilds, David, "Balancing the Corporate Scorecard", In ***The Accounting Review***, v.74, n.4, October 1999, pp. 523-524

Sherman, Daniel J., Olsen, Eugene A., "Stages in the Project Life Cycle in R&D Organizations and the Differing Relationships Between Organizational Climate and Performance", In ***Journal of High Technology Management Research***, v.7, n.1, 1996, pp. 79-90,

Simon, L. P., and Monroe, A. F., "California Provider Group Report Cards: What Do They Tell Us?, In ***American Journal of Medical Quality***, v. 16, n. 2, May-April 2001, pp. 61-69

Sink, D. S., and Tuttle, T. C., Planning and Measurement in Your Organization of the Future, Norcoss GA: Industrial Engineering and Management Press, undated.

Snow, Charles C., Hambrick, Donald C., "Measuring Organizational Strategies: Some Theoretical and Methodological Problems", In ***The Academy of Management Review***, v.5, n.4, October 1980, pp. 527-538

Souder, William, "Award Programs for R&D Personnel", In ***Research-Technology Management***, Nov-December 1985, pp. 13-18

Stainer, Alan, and Nixon, Bill, "Productivity and Performance Measurement in R&D", In ***International Journal of Technology Management***, v.13, n.5/6, 1997, pp. 486-495

Stecher, Brian, An Educator Looks at Quality Improvement in Healthcare, Working Paper, WR-241-EDU, Santa Monica: RAND, February 2005

Stonich, Paul J., "Using Rewards in Implementing Strategy", In ***Strategic Management Journal***, v.2, n.4, October 1981, pp. 345-352

Szonkyi, Robert, "Mechanisms for Improving the Effectiveness of R&D: How Many Mechanisms are Enough?", In ***R&D Management***, v.15, n.3, 1985, pp. 219-225

Szakonyi, R., "Measuring R&D Effectiveness – I", In ***Research-Technology Management***, March-April, 1994 pp.27-32

Szakonyi, R., "Measuring R&D Effectiveness – II", In ***Research-Technology Management***, May-June, 1994b, pp.44-55

Szakonyi, R. (1998), "Leading R&D: How much progress in 10 years?", In ***Research-Technology Management***, Vol. 41, #6, Nov-Dec., 25-29

Takahashi, Tomio, "Management for Enhanced R&D Productivity", In ***International Journal of Technology Management***, v.14, n.6/7/8, 1997, pp. 789-803

Thor, Carl G., "Using a Family of Measures to Assess Organizational Performance", In ***National Productivity Review***, Summer 1995, pp. 111-131

Tipping, J.W., Zeffren, E., and Fusfeld, A.R., (1995), "Assessing the value of your Technology", ***Research-Technology Management***, Sept.-Oct., 1995

Tolchinsky, Paul D., King, Donald C., "Do Goals Mediate the Effects of Incentives on Performance?", In ***The Academy of Management Review***, v.5, n.3, July 1980, pp. 455-467

Tottie, M., Lager, T., "QFD-Linking Customer to the Product Development Process as a part of the TQM Concept", ***In R&D Management***, Vol. 25, #3, 1995, pp.257-267

Towill, Denis R., "Management Theory: Is it of any Practical Use? Or How Does a Fad Become a Paradigm?", In ***Engineering Management Journal***, June 1999, pp. 111-121

Treacy, Michael, Wiersema, Fred, The Discipline of Market Leaders, Perseus Books, 1995

Tuttle, Thomas C., and Sink, D. Scott, "Taking the Threat Out of Productivity Measurement", In ***National Productivity Review***, Winter 1984-85, pp. 24-32

Uchimaru, Kiyoshi, Okamoto, Susumu, Kurahara, Benteru, TQM for Technical Groups: Total Quality Principles for Product Development, Productivity Press, 1993

Vaiana, Mary E., and McGlynn, Elizabeth A., "What Cognitive Science Tells Us About Design of Reports for Consumers", In ***Medical Care Research and Review***, v.59, n.1, March 2002, pp. 3-35

Velury, Jay, "Initiative and Incentive Management: Far From Long Gone", In ***National Productivity Review***, Spring 2000, pp. 73-78

Wagner, John A., "Participation's Effects on Performance and Satisfaction: A Reconsideration of Research Evidence", In ***Academy of Management Review***, v. 19, n.2, 1994, pp. 312-330

Wakim, Nagi T., and Bushnell, David S., "Performance Evaluation in a Campus Environment", In ***National Productivity Review***, Winter 1999, pp. 19-27

Wang, Ching-Wen, Wu, Jue-Ja, Ruey-Yun, Horng, "Creative Thinking Ability, Cognitive Type and R&D Performance", In ***R&D Management***, v.29, n.3, 1999, pp. 247-254

Ward, Anthony, "Measuring the Product Innovation Process", In ***Engineering Management Journal***, October 1996, pp. 242-246

Weller, Ann C., Editorial Peer Review: Its Strengths and Weaknesses, American Society of Information Science and Technology (ASIS&T), 2001

Werner B. M., and Souder, W. E., "Measuring R&D Performance: State of the Art", In ***Research-Technology Management***, Vol. 40, #2, March-April, 1997, pp. 34-42

Werner B. M., and Souder, W. E., "Measuring R&D Performance: US and German Experience", In ***Research-Technology Management***, Vol. 40, #3, May-June, 1997, pp. 28-32

Whetton, David A., and Cameron, Kim S., "Organizational Effectiveness: Old Models and New Constructs", In Jerald Greenberg (Ed.), ***Organizational Behavior: The State of the Science***, Hillsdale, NJ: Eribaum, pp. 135-153

White, Roderick E., "Generic Business Strategies, Organizational Context and Performance: An Empirical Investigation", In ***Strategic Management Journal***, v.7, n.3, May-June 1986, pp. 217-231

Wilkinson, Adrian, Willmott, Hugh, "Total Quality: Asking Critical Questions", In ***The Academy of Management Review***, v.20, n.4, October 1995, pp. 789-791

Seiler, Robert E., Improving the Effectiveness of Research and Development, McGraw-Hill Book Company, 1965

Williamson, Ian O., Cable, Daniel M., "Predicting Early Career Research Productivity: The Case for Management Faculty", In ***Journal of Organizational Behavior***, v. 24, 2003, pp. 25-44

Winn, Joan, "Performance Measures for Corporate Decline and Turnaround", In ***Journal of General Management***, v.19, n.2, Winter 1993, pp. 48-61

Wong, Caroline, Organizational Management of Army Research, RAND Documented Briefing, DB-390-A, 2003

Yin, Robert K., Case Study Research: Design and Methods 2nd Ed., Sage Publications, 1994

Yin, Robert K., Applications of the Case Study Research, Sage Publications, 1993

Yin, Robert K., Case Study Research, Newbury Park: Sage Publications, 1987

Zbaracki, Mark J., "The Rhetoric and Reality of Total Quality Management", In ***Administrative Science Quarterly***, v.43, n.3, Sept. 1998, pp. 602-632

Zucker, Lynne G., Darby, Michael R., Armstrong, Jeff S. "Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology," NBER Working Paper # 8499, National Bureau of Economic Research, 2001