12-2005

Performance Evaluation Models for Strategic Decision-Making: Towards a Hybrid Model

Geraldina Villalobos Quezada
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Education Commons, and the Human Resources Management Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/1052

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
PERFORMANCE EVALUATION MODELS FOR STRATEGIC DECISION-MAKING: TOWARDS A HYBRID MODEL

by

Geraldina Villalobos Quezada

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Educational Studies
Dr. Braaks Applegate, Adviser

Western Michigan University
Kalamazoo, Michigan
December 2005
PERFORMANCE EVALUATION MODELS FOR STRATEGIC DECISION-MAKING: TOWARDS A HYBRID MODEL

Geraldina Villalobos Quezada, Ph.D
Western Michigan University, 2005

Performance management systems serve strategic, administrative, and developmental purposes; therefore, their role in an organization cannot be overestimated. As a function of this strategic role, different evaluation models have been developed and implemented by organizations: BSC, CIPP, TQM, Six Sigma, and AOP.

Following a review of current evaluation theory and practice models that focus on improving strategic decision-making in organizations, four research questions were developed that sought to identify the interrelationships, evaluation components, evaluation indicators, data collected to support the evaluation, evaluation implementation protocols, quantitative and qualitative analyses, outcomes, and critical factors of the BSC and CIPP models.

A multiple case study research method was used to address the study questions. Four BSC and two CIPP cases were studied. A comparison of outcomes revealed that both models were implemented in organizations to maintain focus, assess and monitor performance, reinforce communication of the strategic objectives, and improve performance controls. The BSC model's
implementation protocol followed the five management principles of “Strategy-
Focused Organization.” Alternatively, the CIPP model used four types of
evaluations. Analyses revealed relationships between the BSC and CIPP, such that
both models share compatible evaluation components and collect similar evaluative
information. However, the BSC model cases tended to use quantitative evaluation
indicators, while CIPP cases employed mostly qualitative evaluation indicators.
Both models used tools to develop focus and organizational alignment in their
organizations.

The greatest difference between BSC and CIPP focused on the critical factors
for successful implementation. For BSC, they included management support,
merging it with TQM and Six Sigma, use of software tools, and alignment of
evaluation indicators at all organizational levels. The CIPP model’s critical factors
included stakeholders support, use of different types of evaluation, use of
triangulation methods, and use of communication mechanisms.

Finally, this study proposes a hybrid BSC/CIPP model for strategic decision-
making. Justification for the hybrid model focuses on its value in the context of
managerial strategic decision-making.
ACKNOWLEDGMENTS

The completion of this dissertation represents not just one project seen to completion, but the end of a long and fascinating journey in which many people have played integral parts. Without the support and guidance of so many-and some amazing lucky breaks I might never have made the effort.

I would like to express my gratitude to my very special and brilliant committee members, Dr. Brooks Applegate (my dissertation chair), Dr. Dale Brethower, and Dr. Robert Brinkerhoff. Their patience, guidance and immense knowledge not only shaped me, but also enhanced my skills and abilities in the process. Each of them challenged me with the use of their very own special and unique talents, and in combination, they polished me off and perfected my work, making it worthy for this advanced degree. I was blessed to have three such talented and passionate scholars supporting me, and I am eternally indebted and grateful. Thank you for trusting me, for all your hard work, your sacrifices in time, and the gifts that you all gave me through your involvement. I will cherish them and put them to good use.

Special thanks go to Dr. Van Cooley, Dr. Charles Warfield, and Dr. Jianping Shen for their support throughout all these years at Western Michigan University.

The path of a great education always begins at home. I would like to thank my parents, Geraldina Quesada and Horacio Villalobos, for placing such a high
priority on the quality of my education, and my loved brothers Horacio and
Fernando for giving me all their love and encouragement to take advantage of this
opportunity. Even though my father is not around to see me graduate, I know he is in
spirit. It has long been a dream of mine to honor my parents love and sacrifice with
this accomplishment.

I would certainly not have made it through such a long and arduous process
without my loved uncle Sergio and aunt Peggy Quesada, who encourage me to make
this fascinating journey and have always been very supportive and understanding.
Thank you for all your encouragement, the persistent support you gave me, and
voluminous love that nurtured me and provided all of the impetus needed to embark
upon this journey and to succeed. My Alex, without your love, support and your
unrelenting push, I could not have dreamed or attempted such an enormous
undertaking as this. Thank you for all your encouragement throughout all these
years, and your confidence into my work.

I would also like to express my deepest gratitude to my loving grandmother,
Josefina Aldana, who was very supportive behind the scenes, I am delighted that she
will be around to see me graduate, just after her 92nd. birthday. I would like to thank
also my loving and supportive uncles Emilio and Blanca Quesada, and my
godfather, Ing. Tomas Arocha Morton.

Special thanks go to my loving cousins Sarahita, Mandy, and Sergito,
Emilio, Carlos, and Javier, my sister-in-law Aide, and to all my good friends Amy,
Acknowledgments—Continued

Miki, Adela, Norma, Sylvia and Carlos, Victor and Monica, Marco Antonio Ortega, Smithesh, Ileana, Tim, Hector and Sandra, Jorge and Martha, Lili and Rigo- for their enduring love, support, and encouragement.

This amazing life journey has certainly been all worth it! To all of you...Thank You

Geraldina Villalobos Quezada
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. ii  
LIST OF TABLES ........................................................................ x  
LIST OF FIGURES ........................................................................ xi  

CHAPTER

I. INTRODUCTION ........................................................................... 1  
   Problem Statement .................................................................. 1  
   Background ............................................................................. 5  
   Evaluation Models .................................................................. 5  
      Balanced Scorecard ....................................................... 6  
      CIPP Model ..................................................................... 7  
      Total Quality Management ............................................ 8  
      Six Sigma ....................................................................... 9  
      Anatomy of Performance .............................................. 10  
   Combining BSC, CIPP, TQM, Six Sigma, and AOP Models........... 11  

   Research Questions ............................................................ 14  

II. REVIEW OF LITERATURE ......................................................... 22  
   Introduction .......................................................................... 22  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table of Contents-Continued

CHAPTER

II. REVIEW OF LITERATURE

<table>
<thead>
<tr>
<th>Evaluation Models</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decisions/Accountability Oriented Approaches</td>
<td>22</td>
</tr>
<tr>
<td>Balanced Scorecard (BSC)</td>
<td>25</td>
</tr>
<tr>
<td>Model Overview</td>
<td>25</td>
</tr>
<tr>
<td>Characteristics</td>
<td>27</td>
</tr>
<tr>
<td>Evaluation Components and Evaluation Indicators</td>
<td>29</td>
</tr>
<tr>
<td>used in BSC Model</td>
<td></td>
</tr>
<tr>
<td>BSC and Implementation Protocol</td>
<td>34</td>
</tr>
<tr>
<td>Context, Input, Process, Product (CIPP) Model</td>
<td>37</td>
</tr>
<tr>
<td>Model Overview</td>
<td>37</td>
</tr>
<tr>
<td>Characteristics</td>
<td>39</td>
</tr>
<tr>
<td>Evaluation Components and Evaluation Indicators</td>
<td>43</td>
</tr>
<tr>
<td>used in CIPP Model</td>
<td></td>
</tr>
<tr>
<td>CIPP and Implementation Protocol</td>
<td>49</td>
</tr>
<tr>
<td>Total Quality Management (TQM)</td>
<td>50</td>
</tr>
<tr>
<td>Model Overview</td>
<td>50</td>
</tr>
<tr>
<td>Characteristics</td>
<td>51</td>
</tr>
<tr>
<td>Evaluation Components and Evaluation Indicators</td>
<td>52</td>
</tr>
<tr>
<td>used in TQM Model</td>
<td></td>
</tr>
<tr>
<td>TQM and Implementation Protocol</td>
<td>56</td>
</tr>
</tbody>
</table>
CHAPTER

III. REVIEW OF LITERATURE

Six Sigma .................................................................................... 57
  Model Overview................................................................. 57
  Characteristics............................................................... 58
  Evaluation Components and Evaluation Indicators
  used in Six Sigma Model............................................... 59
  Six Sigma and Implementation Protocol..................... 63
Anatomy of Performance (AOP) .............................................. 65
  Model Overview.................................................................... 65
  Characteristics ................................................................ 66
  Evaluation Components and Evaluation Indicators
  used in AOP Model ....................................................... 70
  AOP and Implementation Protocol ............................. 74
Summary and Comparison of Evaluation Models ................. 77

III. METHODOLOGY ................................................................................. 83

  Case Study Research Methodology ........................................... 84
  Multiple Case Study Design.......................................... 87
  Success Case Method......................................................... 88
  Selection and Description of BSC and CIPP's
  Case Studies ................................................................................. 92
  Sample ......................................................................... 93
CHAPTER

III. METHODOLOGY

Data Preparation ............................................................. 95
Analysis ....................................................................................... 96
Summary...................................................................................... 97

IV. RESULTS ................................................................................... 100
Research Question # 2............................................................... 101
Evaluation Components.......................................................... 102
Evaluation Indicators.............................................................. 104
Data Collected to Support BSC and CIPP Evaluations..... 109
Evaluation Implementation Protocol used in BSC and CIPP Evaluations .............................................................. 111
Qualitative and Quantitative Analyses in BSC and CIPP Evaluations .............................................................. 119
Research Question # 3............................................................... 126
BSC Outcomes....................................................................... 129
CIPP Outcomes...................................................................... 130
Research Question # 4............................................................... 133
BSC Critical Factors............................................................... 139
CIPP Critical Factors............................................................ 142
<table>
<thead>
<tr>
<th>CHAPTER V. DISCUSSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary of Findings and Discussion ..................................</td>
</tr>
<tr>
<td>Decisions/Accountability-Improvement Oriented Evaluation Models and Strategic Decision-Making ..</td>
</tr>
<tr>
<td>BSC and CIPP Evaluation Models ....................................</td>
</tr>
<tr>
<td>Evaluation Components ..............................................</td>
</tr>
<tr>
<td>Evaluation Indicators ...................................................</td>
</tr>
<tr>
<td>Data Collected to Support BSC and CIPP Evaluations ..........</td>
</tr>
<tr>
<td>Implementation Protocol ...............................................</td>
</tr>
<tr>
<td>Qualitative and Quantitative Analyses ........................</td>
</tr>
<tr>
<td>Outcomes of BSC and CIPP .............................................</td>
</tr>
<tr>
<td>Critical Factors of BSC and CIPP .............................</td>
</tr>
<tr>
<td>Evaluation Models and Strategic Decision-Making ..........</td>
</tr>
<tr>
<td>The BSC and CIPP Hybrid Evaluation Model ......</td>
</tr>
<tr>
<td>Limitations .................................................................</td>
</tr>
<tr>
<td>Recommendation ............................................................</td>
</tr>
<tr>
<td>Summary .................................................................</td>
</tr>
</tbody>
</table>

REFERENCES ................................................................................................ 174

APPENDICES

A. Source Information for BSC and CIPP Model Case Studies ........ 181
# LIST OF TABLES

1. The relevance of four evaluation types to decision-making and accountability ................................................................. 42

2. Evaluation indicators used in CIPP model ........................................... 46

3. Methods of potential use in CIPP evaluations ........................................ 48

4. Summary of the AOP results improvement implementation process ..................................................................................... 75

5. Evaluation components used in BSC and CIPP evaluations ............ 102

6. Evaluation indicators used in BSC and CIPP evaluations .............. 105

7. Data collected to support BSC and CIPP evaluations ...................... 110

8. Evaluation implementation protocol used in BSC and CIPP evaluations .............................................................................. 111

9. Qualitative and quantitative analyses in BSC and CIPP evaluations ................................................................................... 119

10. Outcomes from BSC and CIPP evaluations .................................... 127

11. Critical factors from BSC and CIPP evaluations ............................. 133

12. Summary table of findings for BSC and CIPP model characteristics .................................................................................. 156

13. Hybrid model’s components .......................................................... 165
LIST OF FIGURES

1. The BSC major evolutions in applications ............................................ 27
2. The BSC’s implementation process ....................................................... 36
3. Key components of the CIPP evaluation model and associated relationships with programs ............................................................... 41
4. Logic model development ................................................................. 45
5. The Anatomy of Performance (AOP) model ....................................... 69
6. Four views of an organization ............................................................ 72
7. Similarities between BSC and CIPP’s evaluation components ........... 150
8. Hybrid model ..................................................................................... 164
CHAPTER I
INTRODUCTION

Problem Statement

"Organizations are complex enterprises requiring careful leadership to achieve their missions and objectives. In an uncertain environment, characterized by increasing competition for scarce resources, the time allowed to management to make decisions has shortened while the need for timely and meaningful information has increased" (Niven¹, 2003, p.14). As a consequence, accountability and performance measurement have become paramount for organizations. This is illustrated by the following quote, "leaders are dissatisfied with the reliability of traditional measurement tools as organizations are driven to real-time responses. These leaders often feel inundated with data but lacking in relevant performance information-the kind of information that can help make the difference between success and failure" (Klynveld, Peat, Marwick, Goerdeler, 2001, p. 2).

"Measuring and managing performance is a challenging enterprise and seen as one of the keys to managing change and thus gaining competitive advantage in organizations" (Neely, 2004). As Niven (2003) observed "organizations today face increased pressure to implement effective performance management systems and improve operational efficiency, while simultaneously remaining focused on fulfilling their missions" (p. 11).

Performance management systems serve strategic, administrative, and

¹ All references in this dissertation follow APA style as expressed in the American Journal of Evaluation.
developmental purposes (Hayes, Austin, Houmanfar, & Clayton, 2001, p.239); therefore, their role in an organization cannot be overestimated. As a function of this strategic role, different evaluation models have been developed and implemented by organizations (e.g., United Parcel Services (UPS), 1999; Mobil, 2000; Hilton Hotel Corporation, 2000; Spirit of Consuelo, 2002; Nasa, 2004), not only as a means to inform, but additionally to improve both strategic and operational management decision-making. By understanding how these different evaluation models can be used in organizations as strategic management systems, profit and nonprofit organizations can achieve long-term strategic objectives by implementing strategies and linking them to unit and individual goals.

Evaluation is a study designed and conducted to assist some audience to assess an object’s merit and worth (Stufflebeam, 2001). Evaluation models such as the Decision/Accountability-Oriented evaluation approach, (Stufflebeam, Madaus, & Kellaghan, 2000) “provide information needed to both develop and defend a program’s merit and worth by continuously supplying managers with the information they need to plan, direct, control, and report on their programs or spheres of responsibility” (p. 52).

Different evaluation models have also been used to develop and implement strategic performance evaluation that facilitate managers’ strategic decision-making, planning and control. As Norton (2002) observed “the essence of strategy is to define the outcomes that are desired and to identify the assets (tangible and intangible) and activities required to achieve them” (p. 5). Evaluation models also constitute powerful tools for organizational evaluation by providing managers with information
about “what performance is required at the organization, process, and job/performer level, what performance to measure, what questions to ask about performance deviations, and what actions to take to modify performance” (Rummler, 2001, p. 113).

Another use of organizational evaluation models is to help organizations focus not only on traditional performance areas, which tend to look at financial, operational, or functional efficiency, but also focus on non-traditional measures which tend to relate to intangibles such as an entity’s marketplace, stakeholders, strategic implementation, and resource management (Kaplan & Norton, 2004). Ideally, non-traditional measures are usually predictive in nature. Due to information that is focused only on financial measures, organizations have difficulties assessing efficiency and effectiveness. In most cases, when information exists it is limited to financial measures (Kaplan & Norton, 1992). However, according to Kaplan (1996), “Financial measures are inadequate for guiding and evaluating organizations’ trajectories through competitive environments. They are lagging indicators that fail to capture much of the value that has been created or destroyed by managers’ actions in the most recent accounting period. The financial measures tell some, but not all, of the story about past actions and they fail to provide adequate guidance for the actions to be taken today and the day after to create future financial value” (p. 31).

This study was concerned with reviewing the evaluation theory and practice of evaluation models focused on improving strategic and operational management decision-making in organizations. Evaluation models include, the Balanced Scorecard (BSC), (Kaplan & Norton, 1992); the Context, Input, Process, Product (CIPP) model,
(Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985); Total Quality Management (TQM), (Deming, 1920); Six Sigma, (Welch, 1980), Anatomy of Performance (AOP), (Rummler, 2002). These different evaluation models seek to provide necessary information for organizational change, and are used as a means to clarify, communicate and manage strategic decision-making. Addressing the questions of appropriate uses, interrelationships, and outcomes of these evaluation models in practice will provide guidance to evaluators utilizing these models. Finally, a hybrid model was sought that integrated strategic decision-making models. Specifically, this study focused on integrating the BSC and CIPP Models.

The TQM, Six Sigma, and AOP models were chosen, in addition to BSC and CIPP, because they also are performance management and organizational evaluation tools that are commonly applied when BSC and CIPP models are used.

The BSC (Kaplan & Norton, 1992) and CIPP (Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985) models were chosen as the general context of this study because of their comprehensive and similar approach to measuring performance in organizations by facilitating managers' strategic decision-making. A comprehensive literature review of published journal articles failed to identify any studies that explicitly compared BSC with CIPP in the context of their utility for managerial strategic decision-making.

Evaluators and managers need to develop a critical view of the alternatives that can help them consider, assess, and selectively apply optional performance measurement models in order to help them improve their strategic decision-making. Consequently a study of these evaluation models is important as it might help
practitioners identify, examine, and address conceptual and technical issues pertaining to the development and efficient use of these models. A critical review should include, but not be limited to understanding each model in terms of its strengths and weaknesses, determining when and how these models are best applied, developing awareness of how to improve the models, devising better alternatives, and strengthening one's ability to conceptualize hybrid performance evaluation models.

The remainder of this chapter provides (a) the background for this study, (b) the research questions that guided this work, (c) the relevance of this study to the field of evaluation, and (d) definitions. Chapter II contains a review of the literature that focused on the three components of this proposed study: (1) Decision/Accountability-Improvement Oriented Evaluation Models and strategic decision-making, (2) an overview of each evaluation model’s theory including the elements (tools) and interrelationships of these models, and (3) merging BSC with the CIPP Model into a hybrid performance measurement evaluation model. Chapter III outlines the methodology for this study. Chapter IV presents the case studies used that include BSC and CIPP models. Chapter V concludes with a discussion on issues related to the hybrid BSC/CIPP Model and evaluation models’ practices, and presents recommendations for evaluators and researchers.

**Background**

**Evaluation Models**

Managers in organizations are faced with a growing array of evaluation models such as: The BSC (Kaplan & Norton, 1992), the CIPP model (Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985), TQM (Deming, 1920), Six Sigma (Welch, 1980), and AOP (Rummler, 2002), to help them make strategic and operational
management decisions. These evaluation models differ in their orientation, evaluation indicators, information requirements, implementation processes, and outcomes. Analysis of these models is needed in order to provide evaluators and practitioners with an understanding of these distinctions in context/applications, uses, methods, products, strengths, weaknesses, and limitations of each of these evaluation models.

Balanced Scorecard

The BSC was originally developed in the early 1990s by Robert S. Kaplan and David P. Norton in the business/performance measurement area. According to Kaplan and Norton (1996), “A balanced scorecard is a performance measurement system that gives top managers a fast but comprehensive view of their business” (p. 17). Balanced scorecards improve organizational performance by weighting several important measures of organizational performance and linking these to the strategy and vision of the organization. Although companies must adapt balanced scorecard measures to their own vision and strategy, scorecards should portray measures in four different areas: customers, internal processes, financial, and learning and growth.

The BSC is an evaluation model that has been implemented by many organizations (United Parcel Services (UPS), 1999; Mobil, 2000; Hilton Hotel Corporation, 2000; TRURO, 2001; CROSSHOUSE, 2001; Siemens AG, 2004; St. Mary’s Duluth Clinic Health Center, 2004). The BSC has evolved over time into a full Performance Management system applicable to both private sector and public (and not-for-profit) organizations in different areas, such as business (Chang, 2000), manufacturing (Kaplan & Norton, 2004), service (Niven, 2003), and telecommunications (Paladino, 2004). Emphasis has shifted from just the
measurement of financial and non-financial performance, to include the management (and execution) of business strategy in the four different areas.

**CIPP Model**

The CIPP Evaluation Model was developed by Daniel L. Stufflebeam in 1966, and introduced in the education area. The CIPP model has undergone some changes in its application process, from the most fundamental form of CIPP to stress the need for process as well as product evaluations during the first generation, moving into a set of four types of evaluation; context, input, process, and product within a comprehensive system that can be used for summative as well as formative evaluation.

The CIPP Model is a decision-oriented evaluation approach structured to help administrators make good decisions. Under this framework, evaluation is viewed as "the process of delineating, obtaining, and providing useful information for judging decision alternatives" (as cited in Worthen, Sanders, & Fitzpatrick, 1997, p. 154). This evaluation model provides managers and administrators with four different kinds of organizational decisions: Context evaluation, to serve planning decisions, Input evaluation, to serve structuring decisions, Process evaluation, to serve implementing decisions, and Product evaluation, to serve recycling decisions. This comprehensive model is useful for guiding formative and summative evaluations of projects, programs, personnel, products, institutions, and systems. The model has been employed throughout the U.S. and around the world in short-term and long-term investigations (both small and large). Applications have spanned various disciplines and service areas, including education (Horn & McKinley, 2004), housing and
community development (Stufflebeam, 2002), and transportation safety (Stufflebeam & McKee, 2003).

**Total Quality Management**

The early pioneers of quality assurance were Walter Shewhart, Harold Dodge, George Edwards, and others including W. Edwards Deming, who were employees of the Western Electric Company (later Bell Telephone Laboratories) in 1920. These pioneers developed many useful techniques for improving quality and solving quality-related problems. Statistical quality control became widely recognized and the techniques and tools for improving quality developed by these group of pioneers were gradually adopted throughout manufacturing industries. The decade of the 1980s was a period of remarkable change and growing awareness of quality in the U.S. by consumers, industry, and government. As differences in quality between Japanese and U.S. made products were apparent, quality excellence became recognized as a key to worldwide competitiveness and was heavily promoted throughout industry (Evans & Lindsay, 1999, p.7).

“TQM framework is a comprehensive managerial philosophy and a collection of tools and approaches for its implementation. The core principles of total quality are: a focus on the customer, participation and teamwork, and continuous improvement and learning” (Evans & Lindsay, 1999, p. 119). These three principles of total quality are supported and implemented by an integrated organizational infrastructure, a set of management practices, and a wide variety of tools and techniques. “The model has been employed throughout the U.S. and around the world in different sectors, including not only the manufacturing and service sectors

**Six Sigma**

The Six Sigma model originated at Motorola in the early 1980s in response to a CEO-driven challenge to achieve tenfold reduction in product failure levels in five years. In the mid-1990s, Motorola divulged the details of their quality improvement model, which has since been adopted by several large manufacturing companies. In the simplest of terms, Six Sigma is a quality improvement methodology that provides a systematic approach to the elimination of defects that is of importance to the customer. “Six Sigma is a rigorous, focused and highly effective implementation of proven quality principles and techniques. Incorporating elements from the work of many quality pioneers, Six Sigma aims for virtually error free business performance. Sigma, $\sigma$, is a letter in the Greek alphabet used by statisticians to measure the variability in any process. A company’s performance is measured by the sigma level of their business processes” (Pyzdek, 2003, p. 3). Six Sigma tools are applied within the following performance improvement model known as Define-Measure-Analyze-Improve-Control (DMAIC). The tools associated with Six Sigma are qualitative, statistical and instructional devices for observing process variables, quantifying their impact on outcomes, and managing their character (Pyzdek, 2003, p. 4). The six sigma model has been employed in different organizations (General Electric 1987,

Anatomy of Performance

The AOP was developed by Geary Rummler in 2001. The AOP is the theoretical construct or framework underlying an analytical approach that reflects the notion that organizations behave as systems. The AOP framework identifies the major variables impacting individual performance and organization results, and it is based on three principles. First, every organization is a processing and adaptive system. The organization must be aligned. Second, every performer in an organization is in a human performance system. The human performance systems must be aligned. Third, the management system is key to keeping the performance system aligned. Management must be doing the aligning (Rummler, 2002, p. 14).

In order to diagnose where the AOP of a given organization is “broken” or misaligned, leading to sub-par performance, this situation is examined from four views: management, business, performer, and organization system view. From this examination, the root causes of the poor performance in an organization are diagnosed in order to improve and sustain the desired performance. The AOP model has been employed in successful improvement initiatives (Motorola, 1981, U.S. Department of Energy, 2001). “Applications of the AOP model have spanned various service areas, including banking/financial, airline, automotive, telecommunications, hospitality, insurance, manufacturing, healthcare, and pharmaceutical” (As cited in Rummler, 2005).
Combining BSC, CIPP, TQM, Six Sigma, and AOP Models

In an effort to identify where does the CIPP model’s orientation stand in relation to BSC in an evaluation context, Daniel Stufflebeam’s (2001) classification of evaluation approaches is used. Stufflebeam, identified 22 evaluation approaches divided into 4 categories that intend to cover most program evaluation efforts: 2 pseudoevaluations, 13 questions/methods-oriented approaches, 3 improvement/accountability-oriented approaches, and 4 social agenda/advocacy approaches. These evaluation approaches were evaluated against the requirements of the Joint Committee (1994) Program Evaluation Standards, which includes examination of each approach’s utility, feasibility, propriety, accuracy, and overall merit.

The BSC is consistent with the CIPP model in that both of these are Decision/Accountability Oriented approaches intended to provide information to people in organizations to facilitate managers’ strategic decision-making, planning and control. The BSC’s methodology builds on key concepts of evaluation practice that can be found in the CIPP Model (Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985), including customer-defined (i.e., meeting stakeholders needs), continuous improvement, emphasis on organizational effectiveness, and measurement-based management and feedback (Kaplan & Norton, 1992, p. 12). For instance, efforts to improve the quality, responsiveness, and efficiency of internal processes that can be found in the process evaluation foci of the CIPP Model are reflected in the operations portion of the BSC’s internal perspective. Thus, companies already implementing different evaluation models in their initiatives will find ample opportunity to sustain
their programs within the more strategic framework of the Balanced Scorecard. For instance, “the Balanced Scorecard was developed to help managers measure the value of intangible assets such as: skills and knowledge of the workforce, the information technology available to support the workforce, ... The Balanced Scorecard approach has been used to trace the contributions of specific intangibles to specific financial (tangible) outcomes” (Kaplan & Norton, 2004, p. 22).

One parallel between the BSC and TQM, is that both evaluation models place a major consideration in the creation and selection of evaluation indicators. The TQM model’s evaluation indicators should best represent the factors that lead to improved customer, operational, and financial performance. These data and information must be analyzed to support evaluation and decision making at all levels within the company. Thus, a company’s performance and evaluation indicators need to focus on key results (Robin & Kaplan, 1991; Struebing, 1996; English, 1996).

Another parallel between the BSC and TQM evaluation models, is that both of them employ a business performance scorecard. The TQM’s performance scorecard includes a broad set of evaluation indicators that often consists of five key categories: customer satisfaction, financial and market, human resource, supplier and partner performance, and company specific indicators that support the strategy (Evans & Lindsay, 1999, p. 476).

A similarity between the BSC and six sigma evaluation models, is that evaluation indicators of the six sigma model are based on the idea of a balanced scorecard. Balanced scorecards provide the means of assuring that six sigma projects are addressing key business results. Senior management are responsible for
translating the stakeholders' goals into evaluation indicators. These goals and evaluation indicators are then mapped to a strategy for achieving them. Scorecards are developed to display the evaluation indicators under each perspective. Finally, six sigma is used to either close gaps in critical indicators, or to help develop new processes, products, and services consistent with top management's strategy (Pyzdek, 2003, p. 33-34).

Some of the evaluation indicators used in the six sigma model under the four different BSC perspectives, are under the financial perspective (i.e., cost per unit, asset utilization, revenue from new sources, profit per customer), under the customer perspective (i.e., price, time, quality, selection, service relationship), under the internal process perspective (i.e., product introductions revenue, key customer variables, inventory delivery costs, audit results for regulatory compliance), learning and growth perspective (i.e., skills gaps from employee competencies, research deployment time from technology, and employee feedback from corporate culture) (Pyzdek, 2003, p.34).

Concerning, the AOP and BSC models, "it is possible to say that AOP is in agreement with the BSC model regarding the need for managers to have a set of instrument panels to review. Kaplan and Norton call it 'balance.' On the other hand, Rummler calls it tracking the variables that impact the performance of your 'business system.' According to Rummler (2002), these are the following instrument panels that should be tracked in order to examine the variables impacting the performance system: First, external variables, as represented by the external components of the super-system. Second, financial factors. Third, critical success and/or operating
factors (e.g. market share) as determined by the strategy. Fourth, critical resource utilization (e.g., human resources, technology). However, the specific instrument panels and meters in those panels will vary with the organization, based on its strategy and particular industry position (p. 233).

Research Questions

In the context of conducting a decision/accountability evaluation, the following research questions were poised:

1. What are the differences and interrelationships among the BSC, CIPP, TQM, Six Sigma, and AOP evaluation models?

2. What are the similarities and differences related to actual implementation of the BSC and CIPP models in terms of their methods, including: evaluation components, evaluation indicators, data collected to support the evaluation, evaluation implementation protocol, and qualitative and quantitative analyses?

3. What are the outcomes of these two (BSC and CIPP) evaluation models; what are the similarities and differences?

4. What are the critical factors that seem to be associated with successful applications of the BSC and CIPP Model?

Answers to these questions will provide guidance to those evaluators and practitioners interested in implementing and merging different evaluation models, including an understanding of these distinctions in context/applications, uses, methods, products, strengths, weaknesses, limitations, of each of these evaluation models.
Relevance of the Study for Evaluators

Although many different methodologies have been developed and implemented to aid managers to sense earlier and respond more quickly to uncertain changes, managers are still facing challenges when executing strategy. According to Cokins, (2004), “There has been too large a gap between high-end strategy and tactical operational systems to effectively achieve an organization’s strategy, mission, and ultimate vision. In complex and overhead-intensive organizations, where constant redirection to a changing landscape is essential, the linkages between strategy and execution have been coming up short” (p. 12).

Norton (2002) observed that “the essence of strategy is to define the outcomes that are desired and to identify the assets (tangible and intangible) and activities required to achieve them” (p. 5). As already indicated, different evaluation models have been developed and implemented to help address strategic performance evaluation that facilitate managers’ strategic decision-making, planning, and control in organizations.

The different evaluation models represent an array of approaches for examining the impact of an organization’s performance by creating an evaluation system that can be used to inform both strategic and operational management decisions. Various combinations of the different evaluation systems have been designed and implemented these models as a means to address different organizational challenges (United Parcel Services, 1999; Mobil, 2000; Hilton Hotel Corporation, 2000; Siemens AG, 2004; Spirit of Consuelo, 2002; Nasa, 2004). Some of these challenges relate to improving an organization’s measurement system to bring relevant, timely, and
reliable information into the organization’s decision making process, and to aid managers in executing strategy by using them as a means to articulate and communicate strategy, motivate people to execute plans, and to monitor results. The use of these evaluation models aid managers’ decision-making processes by integrating information and developing measures. Together, they impact the organization’s capacity for strategic learning, by providing data that managers can use to determine progress and to make corrective actions that lead to greater effectiveness.

Many research studies (Stufflebeam, Madaus & Kellaghan, 2000; Cokins, 2004; The Balanced Scorecard Collaborative, Hall of Fame Case Studies, Crown Castle, 2004; Shultz, 2004, GE Medical Systems; The Balanced Scorecard Collaborative, Hall of Fame Case Studies, Siemens, 2004) have suggested that these evaluation models can be combined in a hybrid model. A hybrid model may have value because of similar philosophies regarding management, and because may capitalize on different methods of measuring and managing an organization’s performance. Given this, by exploring the uniqueness of each of these evaluation models this study will help evaluators and practitioners to distinguish the unique and complementary values of each model, and to have a better understanding of how the models differ in approach, process and benefits. By understanding what are some of the similarities and differences related to the implementation of the BSC and CIPP Models in terms of the methods that are used in both models may help evaluators to have a broader array of performance evaluation’s tools and methods that can be integrated and applied selectively in performance evaluation contexts. An
understanding and comparison of the different outcomes that can be obtained from these different evaluation models provides an opportunity for evaluators and practitioners to devise better alternatives and solutions to reach the desired outcomes. Finally, a critical review of the critical factors associated with successful applications of both the BSC and CIPP models may help evaluators to understand the strengths and weaknesses of each model, to identify a set of best practices of these models, to understand when and how they are best applied, and to develop an awareness of how to improve the models.

Definitions

*Decision/Accountability-Oriented Evaluation.* The decision/accountability-oriented approach emphasizes that program evaluation should be used proactively to help improve a program as well as retroactively to judge its merit and worth. This approach engages stakeholders as a means to provide focus for the evaluation by addressing their most important questions, providing timely and relevant information to assist decision making, and producing an accountability record. The approach stresses that an evaluation's most important purpose is not to prove, but to improve (Stufflebeam, Madaus, & Kellaghan, 2000, p. 62).

*Evaluation.* The process of determining the merit, worth, or value of some object or the product (i.e., the report) of that process.

*Evaluation Models.* Throughout this dissertation, the term “evaluation models” refers to the performance management and evaluation methodologies that organizations use to inform and to improve both strategic and operational management decisions. Evaluation models discussed in this study, include the
following: The Balanced Scorecard (BSC) (Kaplan & Norton, 1992); the Context, Input, Process, Product (CIPP) model (Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985); Total Quality Management (1972, 1982); Six Sigma (1980), and Anatomy of Human Performance (Rummler, 2002).

Performance Management. The process of managing the execution of an organization's strategy. It addresses the way that plans are translated into results.

Performance Evaluation. The process of assessing program results in terms of established performance indicators.

CIPP Model. The Context, Input, Process, Product Model is a comprehensive framework for guiding formative and summative evaluations of projects, programs, personnel, products, institutions, and systems. The model's core concepts are denoted by the acronym CIPP, which stands for evaluation of an entity's context, inputs, processes, and products. These different types of context, input, process, and product evaluation are typically viewed as separate forms of evaluation; but, they can also be viewed as steps or stages in a comprehensive evaluation.

Balanced Scorecard (BSC). The Balanced Scorecard is a framework to help organizations rapidly implement strategy by translating the vision and strategy into a set of operational objectives that can drive behavior, and therefore, performance. Strategy-driven performance measures provide the essential feedback mechanism required to dynamically adjust and refine the organization's strategy over time. The Balanced Scorecard concept is built upon the premise that what is measured is what motivates organizational stakeholders to act. Ultimately all of the organization's activities, resources, and initiatives should be aligned to the strategy. The Balanced
Scorecard achieves this goal by explicitly defining the cause and effect relationships between objectives, measures, and initiatives across each perspective and down through all levels of the organization (Kaplan & Norton, 2004, p.22).

TQM. “The term total quality management, or TQM, has been commonly used to denote the system of managing for total quality. TQM is a companywide effort, through full involvement of the entire workforce and a focus on continuous improvement that companies use to achieve customer satisfaction. TQM is both a comprehensive managerial philosophy and a collection of tools and approaches for its implementation” (Evans & Lindsay, 1999 p. 118). Total Quality (TQ) “is a people-focused management system that aims at continual increase in customer satisfaction at continually lower real cost. TQ is a total system approach (rather than a separate area or program) and an integral part of high-level strategy; it works horizontally across functions and departments, involves all employees, top to bottom, and extends backwards and forward to include the supply chain and the customer chain. TQ stresses learning and adaptation to continual change as keys to organizational success” (Evans & Lindsay, 1999 p. 118).

Six Sigma. A quality improvement methodology that provides a systematic approach to the elimination of defects that influence something important for the customer (Shultz, 2003). Six Sigma is a rigorous, focused, and highly effective implementation of proven quality principles and techniques. Incorporating elements from the work of many quality pioneers, Six Sigma aims for virtually error free business performance. Sigma, $\sigma$, is a letter in the Greek alphabet used by statisticians.
to measure the variability in any process. A company’s performance is measured by
the sigma level of their business processes” (Pyzdek, 2003, p.3)

*AOP*. The Anatomy of Human Performance is the theoretical construct or
framework underlying an analytical approach that reflects the notion that
organizations behave as systems. The AOP framework identifies the major variables
impacting individual performance and an organization’s results, and it is based on
three principles. First, every organization is a processing and adaptive system. The
organization must be aligned. Second, every performer in an organization is in a
human performance system. The human performance systems must be aligned. Third,
the management system is key to keeping the performance system aligned.
Management must be doing the aligning (Rummler, 2001, p.15).

*Process Evaluation*. In essence, a process evaluation is an ongoing check on a
plan’s implementation plus documentation of the process, including changes in the
plan as well as key omissions and/or poor execution of certain procedures. One goal
is to provide staff and managers feedback about the extent to which staff are
efficiently carrying out planned activities on schedule. Another is to help staff
identify implementation problems and to make needed corrections in the activities or
the plan. Process evaluation information is vital for interpreting product evaluation
results (Stufflebeam, Madaus, & Kellaghan, 2000, p. 294).

*Product Evaluation*. The purpose of a product evaluation is to measure,
interpret, and judge an enterprise’s achievements. Its main goal is to ascertain the
extent to which the evaluand met the needs of all the rightful beneficiaries. A product
evaluation should assess intended and unintended outcomes and positive and negative
outcomes. Product evaluation should also assess long-term outcomes. (Stufflebeam, Madaus, & Kellaghan, 2000, p. 297, 298).

Outcome Evaluation. It is a term applied to activities that are designed primarily to measure the effects or results of programs, rather than their inputs or processes. Outcomes may be related to a target, standard of service, or achievement (Stufflebeam, Madaus, & Kellaghan, 2000, p. 97).
CHAPTER II
REVIEW OF LITERATURE

Introduction

Two central concepts explored in the literature relevant to this dissertation: (1) Decisions/Accountability-Improvement Oriented Evaluation Models and strategic decision-making, (2) an overview of each evaluation model’s theory including the evaluation components, evaluation indicators, data collected to support the evaluation, evaluation implementation protocol, and qualitative and quantitative analyses. The discussion of each concept provides an overview including each evaluation model’s characteristics, evaluation components, evaluation indicators. Some examples of how the BSC and CIPP evaluation models have been implemented and used in organizations are provided in Chapter Four. In conclusion a summary and comparison of evaluation models is discussed.

Evaluation Models
Decisions/Accountability Oriented Approaches

Different evaluation models have been developed and implemented in organizations (United Parcel Services (UPS), 1999; Mobil, 2000; Hilton Hotel Corporation, 2000), such as: The Balanced Scorecard (Kaplan & Norton, 1992); the CIPP model (Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985); Total Quality Management (TQM), (Deming, 1920); Six Sigma, (Welch, 1980), Anatomy of Human Performance (AOP), (Rummler, 2002).

These evaluation models differ in their orientation, information requirements, implementation processes, and outcomes. However, all of these different evaluation
models have a common purpose: they are all used to implement strategic performance evaluation that facilitates managers’ strategic decision-making, planning, and control.

Stufflebeam’s (2001) identified 22 evaluation approaches divided into four categories that intend to cover most program evaluation efforts: two pseudoevaluations, thirteen questions/methods-oriented approaches, three improvement/accountability-oriented approaches, and four social agenda/advocacy approaches.

According to Stufflebeam, Madaus, & Kellaghan (2000) evaluation models that aim to “provide information needed to both develop and defend a program’s merit and worth by continuously supplying managers with the information they need to plan, direct, control, and report on their programs or spheres of responsibility” (p. 52), are categorized as improvement/accountability oriented models. This classification includes also the decisions/accountability oriented approach.

Stufflebeam, Madaus, & Kellaghan (2000) noted, “The decisions/accountability oriented approach emphasizes that program evaluation should be used proactively to help improve a program as well as retroactively to judge its merit and worth. In practice, this approach engages stakeholders in focusing the evaluation, addressing their most important questions, providing timely, relevant information to assist decision making, and producing an accountability record” (p. 62). In this perspective, evaluation models such as: BSC, CIPP, TQM, Six Sigma, and AOP models have commonly sought to address the challenge of providing managers with timely and meaningful information to improve both strategic and operational management decision-making.
Moreover, Stufflebeam, Madaus, & Kellaghan (2000) found that under these circumstances the decision/accountability-oriented approach is useful:

The generic decision situations to be served may include defining goals and priorities, choosing from competing services, planning programs, budgeting, staffing, using services, guiding participation, judging progress, and recycling program operations. Key classes of needed evaluative information are assessments of needs, problems, and opportunities; identification and assessment of competing programs or program approaches; assessment of program plans; assessment of staff qualifications and performance; assessment of program facilities and materials; monitoring and assessment of process; assessment of intended and unintended and short-range and long-range outcomes; and assessment of cost-effectiveness (p. 62).

The intended uses of the different evaluation models mentioned above underline the same decision/accountability oriented approach. For instance, the BSC help managers to formulate and to clarify goals and outcome expectations. The CIPP model not only fosters improvement, but also provides accountability records.

In the evaluation models included in this study the main focus is on improvement, accountability, and enlightenment, which define the purpose of the decision/accountability oriented approach. Stufflebeam, Madaus, & Kellaghan (2000) noted:

“A major advantage of the approach is that it encourages program personnel to use evaluation continuously and systematically to plan
and implement programs that meet beneficiaries' targeted needs. It aids decision making at all program levels and stresses improvement. It also presents a rationale and framework of information for helping program personnel to be accountable for their program decisions and actions. It involves the full range of stakeholders in the evaluation process to ensure that their evaluation needs are well addressed and to encourage and support them to make effective use of evaluation findings. It is comprehensive in attending to context, inputs, process, and outcomes. It balances the use of quantitative and qualitative methods…” (p. 64).

Balanced Scorecard (BSC)

Model Overview

Performance scorecards have a long history of use in organizations (Daniels, 1989; Kaplan & Norton, 1992, Chow, Haddad & Williamson, 1997; Hayes, Austin, Houmanfar, & Clayton, 2001). “The most popular incarnation is likely represented by the recent work of Kaplan and Norton (1992), called the BSC model. The BSC is an evaluation model that weighs several important measures of organizational performance and links these to the strategy and vision of the organization” (Hayes, Austin, Houmanfar, & Clayton, 2001, p.239). “The use of performance and balanced scorecards provides managers with a new evaluation model that includes metrics, such as quality, customer satisfaction, and innovation, that constitute important indicators of business performance that need to be integrated along with financial data” (Kaplan & Norton, 1996, p.6). The word “balance” in the Balanced Scorecard,
represents the balance between financial and non-financial indicators, internal and external constituents of the organization, and lag (generally represents past performance) and lead indicators (performance drivers) (Kaplan & Norton, 1992, p.20).

"The BSC concept is built upon the premise that what is measured is what motivates organizational stakeholders to act. Ultimately all of the organization's activities, resources, and initiatives should be aligned to the strategy. The BSC achieves this goal by explicitly defining the cause and effect relationships between objectives, measures, and initiatives across each perspective and down through all levels of the organization" (Kaplan & Norton, 2004; Niven, 2003; Neely, 1998; Brown, 1996).

Since its development in the early 1990s, the BSC concept and applications have undergone some changes in application. When the BSC concept was developed, it was used as a "tool for performance measurement", and was seen as a method to measure the performance of an organization. The BSC has continued to develop from its most fundamental form as a system for evaluating performance during the first generation of its implementation in organizations, moving into a management system during the second generation, and finally evolving into a universal framework of organizational change in the third generation. Additional elements that are not found in the first and second BSC generations include the use of "strategy maps" to communicate strategy at all levels in the organization. (Morisawa, 2002, p.4). The BSC major evolutions in applications are depicted as follows in Figure 1.
Characteristics

The BSC model’s main characteristics and purposes (Kaplan & Norton, 1992, 2004; Maisel, 1992; Epstein & Manzoni, 1997; Nickols, 2000; Niven, 2003), may be summarized as follows:

An important characteristic of the BSC model is that is used as a valuable evaluation model to enable any person within the organization to pinpoint and track the vital few variables that make or break performance. The BSC model enforces a
discipline around strategy implementation by challenging executives to carefully translate their strategies into objectives, measures, targets, and initiatives in four balanced perspectives: customer, financial, internal processes, and learning and growth. Another characteristic of the BSC model is that it facilitates managers’ strategic decision-making, planning and control in organizations, by aiding people to think in terms of syndrome dynamics and connectivity. The BSC is an important tool that captures hypotheses of strategy and enables measurement development. Additionally, the BSC model serves strategic purposes, by employing it as the foundation of an integrated and iterative strategic management system. Organizations are using the BSC model to:

- Clarify and update strategy
- Communicate strategy throughout the company
- Align unit and individual goals with the strategy
- Link strategic objectives to long-term targets and annual budgets
- Identify and align strategic initiatives
- Conduct periodic performance reviews to learn about and to improve strategy

The BSC model enables a company to align its management processes and focuses the entire organization on implementing long-term strategy. One of the main purposes of using balanced scorecards in organizations is to drive the process of change. By feeding systems for organizational learning, where managers have quantified measures that let them make “fact-based” decisions about where they must change to successfully execute the strategy and continue to add value to the organization over the long term. The BSC is also a valuable tool for accountability.
purposes, and broadens and deepens relationships with stakeholders. Today, to secure
the loyalty of increasingly powerful customers, employees, and shareholders,
managers need to develop and report measures that demonstrate that the organization
is delivering the value demanded.

**Evaluation Components and Evaluation Indicators used in BSC Model**

As mentioned above, although companies must adapt balanced scorecards
measures to their own vision and strategy, scorecards should portray measures in four
different areas (Kaplan & Norton, 1992, 1996; Chow, Haddad & Williamson, 1997;
Niven, 2003):

- **Customer Perspective.** The balanced scorecard demands that managers
translate their general mission statement on customer service into specific measures
that reflect the factors that really matter to customers. Customers' concerns tend to
fall into four categories: time, quality, performance and service, and cost.

- **Internal Business Perspective.** Customer-based measures must be translated
into measures of what the company must do internally to meet its customers’
expectations. The internal measures should stem from the business processes that
have the greatest impact on customer satisfaction; one example is the factors that
affect cycle time, quality, employee skills, and productivity. Companies should also
attempt to identify and measure their company’s core competencies, which are the
critical technologies needed to ensure continued market leadership. Companies
should identify the processes and competencies at which they must excel and specify
measures for each. It is also important to mention, that in order for companies to
achieve goals on cycle time, quality, cost, partnering, and marketing, managers must
devise measures that are influenced by employees’ actions. In this way, employees at lower levels in the organization have clear targets for actions, decisions, and improvement activities that will contribute to the company’s overall mission.

**Innovation and Learning Perspective.** A company’s ability to innovate, to improve, and to learn ties directly to the company’s value. Only through the ability to launch new products, create more value for customers, and continually improve operating efficiencies can a company penetrate new markets and increase revenues and margins, and therefore grow and increase shareholder value.

**Financial Perspective.** Financial performance measures indicate whether the company’s strategy, implementation, and execution are contributing to bottom-line improvement. Typical financial goals have to do with profitability, growth, and shareholder value. However, some of the problems resulting when managers focus only on financial measures are the backward-looking focus, and their inability to reflect contemporary value-creating actions.

Some of the elements and tools used in the BSC model are as follows: strategy maps, measures, targets, and initiatives. These BSC elements can be linked. For example: business strategies give managers the approach chosen to meet customer needs and attain the desired goals. Strategies are made up of building blocks that can be mapped and measured with performance measures. Targets give managers the expected levels of performance that are desired. New initiatives provide new information to successfully meet challenges and test strategy assumptions (Kaplan & Norton, 1996; Neely, 2004; Niven, 2005).
A measure, is a statement of how success in achieving an objective will be measured and tracked. Measures are written statements of what we will track and trend over time, not the actual targets such as direction and speed. A measure should include a statement of the unit to be measured (\$, headcount, \%, rating). Examples include: "Year over Year Sales (\$)" (Financial), "Customer Satisfaction Rating" (Customer), "Service Error Rate (\%)" (Internal), "Strategic Skills Coverage Ratio" (Learning & Growth) (Kaplan & Norton, 1996; Cokins, 2004; Niven, 2005).

A target is the level of performance or rate of improvement required for a particular measure. Targets are stated in specific units (\$, #, \%, rating, etc.), and should include time-based segments (annually, quarterly, etc.) as appropriate. Targets should be observed over time to determine important trending behavior so that corrective action can be taken as needed (Kaplan & Norton, 1996, Cokins, 2004; Niven, 2005).

An initiative is a key action program developed to achieve objectives or close the gap between measures, performance, and targets. Initiatives are often known as projects, actions, or activities. They differ from objectives in that they are more specific, have stated boundaries (beginning and end), have a person/team assigned to accomplish them, and have a budget. Several initiatives taken together may support a specific objective or theme. It is important for an organization to define the boundaries for initiatives, such as "all strategic projects over \$500k in size". It is also important that initiatives be strategic in nature, and not "operations as usual", such as "Recruit a new Sales Rep." Examples include: "Develop Quality Management..."

The BSC model uses strategy maps as a visual representation of an organization's strategy and the processes and systems necessary to implement that strategy (Cokins, 2004). “A strategy map will show employees how their jobs are linked to the organization's overall objectives. The strategy map is used to develop the Balanced Scorecard. Themes are one of the major components of an organization's strategy, providing an overview of how an organization will reach its strategic destination (or five-year plan). An organization's destination can usually be broken down into three or four basic themes that may cross all perspectives. Themes are the pillars of a Strategy Map” (Kaplan & Norton, 2004, p. 30).

The following paragraph describes the process of defining the evaluation indicators, data collected to support the evaluation of BSC, implementation protocol, and qualitative and quantitative analyses employed in the BSC model:

In terms of what is measured, a BSC evaluation views the organization from four different perspectives (customer, internal business, innovation and learning, and financial perspective). Then for each objective there are metrics (in evaluation practice BSC metrics are called ‘evaluation indicators) relative to each of these perspectives. Thus, evaluation indicators must be developed based on the priorities of the strategic plan, which provides the key business drivers and criteria that metrics managers most desire to watch. Processes are then designed to collect information relevant to these evaluation indicators and reduce it to numerical form for storage, display, and analysis. Decision makers examine the outcomes of various measured
processes and strategies and track the results to guide the company and provide feedback (Cokins, 2004). The BSC evaluation indicators on each of the different perspectives become the standards used to evaluate and communicate performance against expected results. In evaluation practice, standards are the criteria to evaluate those outcomes or indicators that were agreed upon with decision makers at the beginning of the evaluation process.

The BSC evaluation indicators must be derived from the company's strategy; and provide critical data and information about key processes, outputs, and results. Data and information needed for BSC implementation are of many types, including customer, product and service performance, operations, market, competitive comparisons, supplier, employee-related, and cost and financial. Analysis entails using data to determine trends, projections, and cause and effect, which might not be evident without analysis. Data and analysis support a variety of company purposes, such as planning, reviewing company performance, improving operations, and comparing company performance with competitors' or with best practices' benchmarks. The BSC evaluation indicators are measurable characteristics of products, services, processes, and operations the company uses to track and improve performance. The measures or indicators should be selected to best represent the factors that lead to improved customer, operational, and financial performance. A comprehensive set evaluation indicators tied to customer and/or company performance requirements represents a clear basis for aligning all activities with the company's goals. Through the analysis of data from the tracking processes, the
measures or indicators themselves may be evaluated and changed to better support such goals. (Kaplan & Norton, 2004; Cokins, 2004; Niven, 2005).

Different methods and statistical tools (e.g., outlier detection, regression analysis, data mining, strategy maps, clustering methods) and certified software (e.g., SAS, Hyperion, CorVu, Bitam) are used in the BSC. Some of them include: interviews, observations, case studies, checklists, focus groups, annual reports to shareholders, strategic plan, operational plan, monthly performance reports reviewed by senior executives, finance data, marketing/customer service data, human resource data, competitor data, industry studies, consultant studies, by comparing outcomes to goals and targets set for each measure.

Both qualitative and quantitative analyses is used to report on BSC evaluation indicators. The BSC may also use robust statistical tools to measure and manage data in organizations. For instance: BSC software applications are instrumental in collecting, analyzing performance, and communicating performance information. BSC certified software (Cokins, 2004) enables organizations to implement the BSC organization wide, to see the causes and effects of an organization's strategy, and to identify sources of business failure and isolate BSC best practices that lead to success.

**BSC and Implementation Protocol**

In general, the BSC’s implementation process entails four different stages (Kaplan & Norton, 1996):

First, clarify and translate vision and strategy. In this stage, the senior executive management team works together to translate its business unit's strategy into specific strategic objectives. Financial and customer objectives are set first,
emphasizing aspects such as revenue and market growth, profitability, and customer and market segments in which the company has decided to compete. With financial and customer objectives established, an organization then identifies the objectives and measures for its internal business process. The BSC highlights those key processes for achieving breakthrough performance for customers and shareholders. Learning and growth objectives are also identified and involve investments in people, systems, and procedures, such as training employees, information technology and systems, and enhanced organizational procedures.

Second, communicate and link strategic objectives and measures. The goal in this stage is to communicate the BSC strategic objectives and measures throughout the organization using different means. Such means include, company newsletters, bulletin boards, videos, and networked personal computers. Everyone in the organization should be able to understand the business unit’s long-term goals, as well as the strategy for achieving these goals. All organizational efforts and initiatives should be aligned then to the needed change processes.

Third, plan, set targets, and align strategic initiatives. In the third stage, senior executives should establish targets for the scorecard measures, three to five years out. Executives then should identify stretch targets for an organizations’ financial, customer, internal-business process, and learning and growth objectives. These stretch targets can come from several sources. For instance, benchmarking can be used to incorporate best practices. Once these targets are established, managers can align their strategic quality, response time, and reengineering initiatives for achieving the objectives. Moreover, Kaplan and Norton (1996) observed that the planning and
target-setting management process enables the organization to: (a) quantify the long-term outcomes it wishes to achieve; (b) identify mechanisms and provide resources for achieving those outcomes; and (c) establish short-term milestones for the financial and non-financial measures on the scorecard.

Fourth, enhance strategic feedback and learning. This final stage is considered by Kaplan and Norton to be the most innovative and most important aspect of the entire scorecard management process. This process provides the capability for organizational learning at the executive level. Managers are then provided with a procedure to receive feedback about their strategy and to test the hypotheses on which the strategy is based. The BSC enables them to monitor and adjust the implementation of their strategy and, if necessary, to make fundamental changes in the strategy itself (p.10-12).

Figure 2. The BSC's implementation process

The BSC’s implementation process undergo three sequential phases (Norton, 2002): the first phase “mobilization” includes a three-to six-month period that was devoted to executive-level momentum building by communicating the need for change, building the leadership team, and clarifying the vision/strategy. Balanced scorecards help clarify the strategy. The use of the customer as a focal point in the new strategies plays an important role in the change process. Finally, organizations need to develop a leadership team to help them guide the process of change. The second phase was related to the design and rollout of the BSC and incorporated a six-month period in which the new strategy was rolled out at the top levels of the organization. Balanced Scorecards were used to cascade, link, and align this rollout process. The final phase of sustainable execution included a 12-to 24-month period where the strategy was integrated into the day-to-day work and culture of the organization.

Context, Input, Process, Product (CIPP) Model

Model Overview

The CIPP evaluation model was developed by Daniel L. Stufflebeam, and introduced in 1966 in the education area. The CIPP evaluation model has been an influential proponent of a decision-oriented evaluation approach structured to help administrators make good decisions. The CIPP evaluation model serves managers and administrators to face four different kinds of organizational decisions: Context evaluation, to serve planning decisions; Input evaluation, to serve structuring decisions; Process evaluation, to serve implementing decisions; and Product evaluation, to serve recycling decisions. Furthermore, the CIPP is a comprehensive
model for guiding formative and summative evaluations of projects, programs, personnel, products, institutions, and systems. The model has been employed throughout the U.S. and around the world in short-term and long-term investigations (both small and large). Applications have spanned various disciplines and service areas, including education (Horn & McKinley, 2004), housing and community development (Stufflebeam, 2002), and transportation safety (Stufflebeam & McKee, 2003). “The CIPP model emphasizes that evaluation’s most important purpose is not to prove, but to improve” (Hanssen, 2004, p. 14).

The CIPP model has undergone some changes in its application process, from the most fundamental form during the first generation of CIPP (Stufflebeam, 1966), that stressed the need for process as well as product evaluations. The second generation published a year later (Stufflebeam, 1967) included context, input, process, and product evaluations and emphasized that goal-setting should be guided by context evaluation, including a needs assessment. Additionally, it emphasized that program planning should be guided by input evaluation, including assessments of alternative program strategies. The third generation (Stufflebeam, D.L. Foley, W.J., Guba, E.G., Hammond, R.L., Merriman, H.O., & Provus, M., 1971) set the four types of evaluation within a systems, improvement-oriented framework. The fourth generation (Stufflebeam, 1972), showed how the model could and should be used for summative as well as formative evaluation. Finally, the model fifth’s generation (Stufflebeam, 2002) breaks out product evaluation into the four types of evaluation to help assess a program’s long-term viability.
Characteristics

Corresponding to the letters in the acronym CIPP, this model's core parts are context, input, process, and product evaluation. In general, these four parts of an evaluation respectively ask, "What needs to be done? How should it be done? Is it being done? and Did it succeed?" (Stufflebeam & Mc Kee, 2003):

**Context Evaluation.** Is a type of evaluation that serves to plan decisions. This type of evaluation helps managers to determine what needs are to be addressed by a program and to define objectives for the program. Context evaluation asks, "What stakeholder's needs should be addressed?" (p. 2).

**Input Evaluation.** Is a type of evaluation that serves to structure decisions. This type of evaluation helps managers to determine what resources are available, what alternative strategies for the program should be considered, and what plan seems to have the best potential for meeting needs facilitates design of program procedures. Input evaluation asks, "What facilities, materials, and equipment are needed?" (p. 2).

**Process Evaluation.** Is a type of evaluation that serves to implement decision making. Process evaluation asks, "How well is the plan being implemented? What barriers threaten its success? What revisions are needed?" Once these questions are answered, procedures can be monitored, controlled, and refined (p. 2).

**Product Evaluation.** Is a type of evaluation that serves to recycle decisions. Product evaluation asks, "What results were obtained? How well were needs reduced? What should be done with the program after it has run its course?" These questions are important in judging program attainments (p. 2).
Figure 3, summarizes the CIPP model's basic elements in three concentric circles and portrays the central importance of defined values (Stufflebeam & Mc Kee, 2003):

The inner circle denotes the core values that should be identified and used to ground a given evaluation. The wheel surrounding the values is divided into four evaluative foci associated with any program or other endeavor. The four foci are goals, plans, actions, and outcomes. The outer wheel indicates the type of evaluation that serves each of the four evaluative foci. These types of evaluation include, for instance, context, input, process, and product evaluation. Each two-directional arrow represents a reciprocal relationship between a particular evaluative focus and a type of evaluation. The goal-setting task raises questions for a context evaluation, which in turn provides information for validating or improving goals. Planning improvement efforts generate questions for an input evaluation, which correspondingly provide judgments and direction for strengthening plans. Program actions bring up questions for a process evaluation, which in turn provides judgment of activities and feedback for strengthening staff performance. Product evaluations focus on accomplishments, lack of accomplishments, and side effects, in order to judge program outcomes and to identify needs for achieving better results (p. 5).
Figure 3. Key components of the CIPP evaluation model and associated relationships with programs


According to Stufflebeam, 2003; Candoli, Cullen & Stufflebeam, 1997; Finn et. al, 1997) the CIPP model has been used by evaluators as a useful guide for decision-making improvement and accountability purposes from a formative and summative orientation as shown in Table 1.
### Table 1.

**The relevance of four evaluation types to decision-making and accountability**

<table>
<thead>
<tr>
<th>Formative Evaluation:</th>
<th>Summative Evaluation:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prospective application of CIPP information to assist decision-making and quality assurance</td>
<td>Retrospective use of CIPP information to sum up the program's merit, worth, probity, and significance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context</th>
<th>Guidance for identifying needed interventions and choosing and ranking goals (based on assessing needs, problems, assets, and opportunities).</th>
<th>Comparison of goals and priorities to assessed needs, problems, assets, and Opportunities.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>Guidance for choosing a program or other strategy (based on assessing alternative strategies and resource allocation plans) followed by examination of the work plan.</td>
<td>Comparison of the program’s strategy, design, and budget to those of critical competitors and to the targeted needs of beneficiaries.</td>
</tr>
<tr>
<td>Process</td>
<td>Guidance for implementing the work plan (based on monitoring and judging activities and periodic evaluative feedback).</td>
<td>Full description of the actual process and record of costs. Comparison of the designed and actual processes and costs.</td>
</tr>
<tr>
<td>Product</td>
<td>Guidance for continuing, modifying, adopting or terminating the effort (based on assessing outcomes and side effects).</td>
<td>Comparison of outcomes and side effects to targeted needs a and, as feasible, to results of competitive programs.</td>
</tr>
</tbody>
</table>

Two primary purposes are found in the CIPP model as a decision/accountability approach, and are as follows (Stufflebeam, 2001, p.56): First, to provide knowledge and a value base for making and being accountable for decisions that result in developing, delivering, and making informed use of cost-effective services. Second, to judge alternatives for defining goals and priorities, choosing from competing services, planning programs, budgeting, staffing, using services, guiding participation, judging progress, and recycling program operations.

**Evaluation Components and Evaluation Indicators used in CIPP Model**

The essential evaluation components are each type of evaluation (context, input, process, and product) of the model including:

**Context evaluation.** This type of evaluation is employed to assess needs, problems, assets, and opportunities within a defined environment. The following four elements are critically important in designing a sound context evaluation program or project: First, identification of client’s needs in order to accomplish the program’s goals and objectives. Second, recognition of problems that need to be addressed in order to meet targeted needs. Third, examination of resources that should include accessible expertise and services to help fulfill the targeted purpose. Fourth, identification of opportunities to support efforts the evaluation efforts and to meet needs and solve associated problems.

**Input evaluation.** This type of evaluation is used to assess the proposed program, project, or service strategy including the work plan and budget for carrying out the effort. Additionally, it assist managers by identifying, examining, and carrying
out those potentially relevant approaches and assess the clients' business environment for political barriers, financial, or legal constraints, and potential resources.

**Process evaluation.** This type of evaluation is used to assess and strengthen a program's implementation process. Process evaluation help managers to document the implementation process, so that they can obtain feedback about the extent to which staff are carrying out planned activities on schedule, as planned, and efficiently. Additionally, it help managers to identify implementation problems and to make needed corrections in the activities.

**Product evaluation.** This type of evaluation is used to assess a program's intended and unintended and positive and negative outcomes. Its main purpose is to determine the extent to which the program or project met the needs of the client. Product evaluation is sometimes divided into impact, effectiveness, sustainability, and transportability evaluation components to assess long-term outcomes. A product evaluation should assess those outcomes obtained at the team and individual levels.

The CIPP model might use logic models in some evaluations (Stufflebeam, 1995; Coffman, 1999) not only to display the inter-relationship of goals and objectives (the emphasis is on short-term objectives as a way to achieve long-term goals), but also to link the various activities together in a manner that indicates the process of program implementation.

"Logic models are also used to find gaps in the program theory and work to resolve them, focus the evaluation around essential linkages of "questions," engage the stakeholders in the evaluation, and build a common sense of what the program is
all about and how the parts work together” (W.K. Kellogg Foundation, 2000, p. 5).

An illustration of a logic model development is provided in Figure 4.

*Figure 4. Logic model development*

<table>
<thead>
<tr>
<th>RESOURCES</th>
<th>ACTIVITIES</th>
<th>OUTPUTS-SHORT</th>
<th>SHORT-&amp;LONG-TERM OUTCOMES</th>
<th>IMPACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>In order to accomplish our set of activities we will need the following:</td>
<td>In order to address our problem or asset we will accomplish the following activities:</td>
<td>We expect that once accomplished these activities will produce the following evidence or service delivery:</td>
<td>We expect that if accomplished these activities will lead to the following changes in 1-3 then 4-6 years:</td>
<td>We expect that if accomplished these activities will lead to the following changes in 7-10 years:</td>
</tr>
</tbody>
</table>


In terms of what is measured, the CIPP model evaluates an organizations’ programs from the above mentioned four types of evaluation (context, input, process, and product). Then for each type of evaluation there are evaluation indicators relative to each of these. Thus, evaluation indicators in the CIPP model must be developed based on the goals and objectives of the evaluation. As noted earlier, in order to determine which parts of the CIPP model to employ and what information to collect, the evaluator needs to identify and address the client’s purpose for the evaluation, which provides the key criteria for the indicators to include in the evaluation.

The CIPP evaluation indicators on each of the different model’s core parts (context, input, process, and product) become the standards used to evaluate and
communicate performance against expected results. For instance Table 2 provides some examples of evaluation indicators that are used under the four CIPP model’s core parts.

Table 2.

**Evaluation indicators used in CIPP model**

<table>
<thead>
<tr>
<th>Context</th>
<th>Input</th>
<th>Process</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Life:</td>
<td>Planning:</td>
<td>Supervision:</td>
<td>Impact Evaluation:</td>
</tr>
<tr>
<td>Health</td>
<td>Values Clarification</td>
<td>Scheduling</td>
<td>Percent of target group served</td>
</tr>
<tr>
<td>Education/training</td>
<td>Defined target group</td>
<td>Implementing plans</td>
<td>Levels of participation</td>
</tr>
<tr>
<td>Social</td>
<td>Clear goals</td>
<td>Progress objectives</td>
<td>Effects on the community</td>
</tr>
<tr>
<td>Community Setting:</td>
<td>Preparation:</td>
<td>Resource Mgmt:</td>
<td>Effectiveness Evaluation:</td>
</tr>
<tr>
<td>Resource organization</td>
<td>Commitment resources</td>
<td>Fiscal records</td>
<td>Full range of outcomes</td>
</tr>
<tr>
<td>Government services</td>
<td>Budget</td>
<td>Resource utilization</td>
<td>Depth of effects</td>
</tr>
<tr>
<td>Economy</td>
<td>Training &amp; Evaluating</td>
<td>Cost overruns</td>
<td>Short-term outcomes</td>
</tr>
<tr>
<td>Political Climate</td>
<td>Facilities</td>
<td>Quality Control</td>
<td>Long-term outcomes</td>
</tr>
<tr>
<td>Related Programs</td>
<td>Equipment</td>
<td>Internal evaluation</td>
<td>Unintended outcomes</td>
</tr>
<tr>
<td>Employment</td>
<td>Safety standards</td>
<td>Correction of-</td>
<td>Cost-effectiveness</td>
</tr>
<tr>
<td>Private sector leaders</td>
<td>Publicity</td>
<td>operational problems</td>
<td>Sustainability</td>
</tr>
<tr>
<td>Recreation opportunities</td>
<td>Policy decisions</td>
<td>External relations</td>
<td>Institutionalization plans &amp; Work schedule</td>
</tr>
</tbody>
</table>

A reporting plan written by the evaluator is employed to promote the use of findings in CIPP evaluations. "This report should involve clients and other audiences (especially targeted users) in deciding the content, nature, and timing of needed reports. The evaluators should engage the client and other intended users to help in planning how the evaluator will disseminate findings. Means for disseminating findings include, oral reports, hearings, community forums, focus groups to examine and respond to findings, multiple reports targeted to specified audiences, press releases, sociodramas to portray and explore the findings, and feedback workshops aimed at applying the findings" (Stufflebeam, Madaus, & Kellaghan, 2000).

"The CIPP Model uses multiple qualitative and quantitative methods, and triangulation procedures to assess and interpret a multiplicity of information" (Stufflebeam, 2003; Horn, 2004). These different methods applied in the CIPP model are used in context, input, process, and product types of evaluation. The use of multiple methods for each type of evaluation provides needed crosschecks on findings. According to Denzin (1978), "whereby a variety of data sources, different perspectives or theories, different methods, and even different investigators are pitted against one another in order to cross-check data and interpretation." Additionally, depending on the program or project evaluation’s purpose qualitative and quantitative methods may be combined to strengthen the evaluation results. An illustration of the various methods of potential use in CIPP evaluations is provided in Table 3.
### Table 3.

*Methods of potential use in CIPP evaluations*

<table>
<thead>
<tr>
<th>Methods</th>
<th>Transportability</th>
<th>Context</th>
<th>Input</th>
<th>Process</th>
<th>Impact</th>
<th>Effectiveness</th>
<th>Sustainability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Literature Review</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Document Review</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Visits to Other Programs</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Advocate Teams (to create &amp; assess competing action plans)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delphi Technique</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Program Profile/Database</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Case Studies</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparative/Experimental Design Studies</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Stakeholder Interviews</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Focus Groups</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cost Analysis</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Secondary Data Analysis</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal-Free Evaluation</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Photographic Record</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Task Reports/Feedback Meetings</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Synthesis/Final Report</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Evaluative information that is important to include under the CIPP model is as follows: First, a thorough assessment of needs, problems, and opportunities. Second, an identification of similar programs or approaches. Third, a review of program plans and staff competencies. Fourth, an identification of program facilities and resources. Fifth, a continuous monitoring of process is needed. Sixth, an assessment of intended and unintended and short-range and long-range outcomes. Seventh, a calculation of the return on investment ratio of the cost and benefits obtained from the implementation of the program is important.

**CIPP and Implementation Protocol**

The CIPP model is a flexible evaluation model that provides managers with the opportunity to choose the type(s) of evaluation (context, input, process, and product) that are needed to conduct in order to conduct an evaluation of a program and meet the identified client’s needs. In order to determine which parts of the CIPP model to employ and what information to collect, the evaluator needs to identify and address the client’s purpose for the evaluation. Additionally, the CIPP model includes a summative and formative evaluation components. A summative evaluation includes all four types of evaluation in order to describe the program, whereas a formative evaluation might just focus on only the type(s) of evaluation needed to guide certain program decisions or to answer specific evaluation questions. Moreover, in assessing context, input, process, and product the evaluator should compile the information required by each pertinent type of evaluation.
Once the evaluator and the client have identified the purpose of the evaluation, and which parts of the CIPP model to employ, and what information to collect, the evaluator needs to design the work that needs to be done.

**Total Quality Management (TQM)**

*Model Overview*

The term total quality management, or TQM, has been commonly used to denote the system of managing for total quality (Evans & Lindsay, 1999). The TQM evaluation model is a companywide effort through full involvement of the entire workforce that focuses on continuous improvement that companies use to achieve customer satisfaction (Reimann, 1989; Schmidt & Finnigan, 1992, Hunt, 1993, Evans & Lindsay, 1999; Pyzdek, 2003). “TQM framework is a comprehensive managerial philosophy and a collection of tools and approaches for its implementation. The core principles of total quality are: a focus on the customer, participation and teamwork, and continuous improvement and learning” (Evans & Lindsay, 1999, p. 119). These three principles of total quality are supported and implemented by an integrated organizational infrastructure, a set of management practices, and a wide variety of tools and techniques.

The TQM model has been employed throughout the U.S. and around the world in different sectors, including not only the manufacturing and service sectors (Milliken & Company, 1989; AT&T, 1992), but also marketing and sales (Ames Rubber Corporation, 1993), product design and engineering (Motorola, 1988), purchasing and receiving (Wallace Company, Inc, 1990), finance and accounting
Since its development in the early 1900s, the TQM concepts and its applications have undergone some major evolutions. The early pioneers of quality assurance were W. Edward Deming, Walter Shewhart, Harold Dodge, George Edwards and others, who were employees of the Western Electric Company (later Bell Telephone Laboratories) in the 1920s. These pioneers developed many useful techniques for improving quality and solving quality problems. Statistical quality control became widely known and the techniques and tools for improving quality developed by this group of pioneers were gradually adopted throughout manufacturing industries. The decade of the 1980s was a period of remarkable change and growing awareness of quality in the United States by consumers, industry, and government. During this time, the differences in quality between Japanese and United States made products were apparent and quality excellence became recognized as a key to worldwide competitiveness and was heavily promoted throughout industry (Evans & Lindsay, 1999).

**Characteristics**

“Total Quality (TQ) is a people-focused management system that aims at continual increase in customer satisfaction at continually lower real cost. TQ is a total system approach (not a separate area or program) and an integral part of high-level strategy; it works horizontally across functions and departments, involves all employees, top to bottom, and extends backwards and forward to include the supply chain and the customer chain. TQ stresses learning and adaptation to continual
change as keys to organizational success. The core principles of total quality include a focus on the customer, participation and teamwork, and continuous improvement and learning” (Evans & Lindsay, 1999 p. 118-119).

**Evaluation Components and Evaluation Indicators used in TQM model**

The three principles of total quality are supported and implemented by an integrated organizational infrastructure, a set of management practices, and a wide variety of tools and techniques (Evans & Lindsay, 1999):

**Infrastructure.** This component refers to the fundamental management systems that need to be in place for successful organizational performance, and includes the following elements:

**Leadership.** Under this component managers should commit to act as change agents for quality. Some of the fundamental questions that managers in an organization should address are: How does managers are creating and sustaining values, setting company directions, developing and improving an effective leadership system?

**Strategic business planning.** This component constitutes the driver for quality improvement throughout the organization. Under this element, some of the fundamental evaluation questions that managers in an organization should address are: Who are our customers? What is our mission? What principles do we value? What are our long-range and short-range goals? How do we accomplish these goals?

**Human resources management.** Under this component employees should align their work to meet the company’s quality and performance goals. This can only be achieved through appropriate employees’ education and training. Under this element,
some of the fundamental questions that managers in an organization should address are: How are managers designing work and jobs that encourages all employees to contribute effectively to achieve the organization’s performance and learning objectives? How are managers designing compensation and recognition systems to reinforce performance?

Process management. Under this component processes are developed in order to create value for customers. This process management perspective aims to provide employees with a holistic picture of the different parts of the organization in order to help them understand how the organization works as a total system. In addition, it helps managers to recognize that problems arise from processes, not people. Under this element, some of the fundamental questions that managers should address are: How does the organization design products, services, and production delivery processes to incorporate changing customer requirements, meet quality and operational performance requirements, and ensure trouble-free introduction and delivery of products and services?

Data and information management. Under this component evaluation indicators are derived from the organization’s strategy and provide critical data and information to managers about key processes, products, services, and results. Some of the fundamental questions that managers should address are: How are managers selecting, managing, using information and data to support key company processes and improve an organization’s performance? How are managers reviewing the organization’s performance and capabilities to assess progress and determine improvement priorities?.
Many types of data and information are needed for quality assessment and quality improvement, including customer needs (Hayes, 1990; Rosenberg, 1996) product and service performance (Berry, Valarie, & Parasuraman, 1990) operations performance (Haywood, 1988) market assessments (Goodman, DePalma, & Breetzmann, 1996), supplier performance (Lovitt, 1989; Stundza, 1991), and employee performance (Williams, 1995; Ingle, 1982). One parallel between the BSC and TQM, is that both evaluation models place a major consideration in the creation and selection of evaluation indicators. The TQM model's evaluation indicators should best represent the factors that lead to improved customer, operational, and financial performance. These data and information must be analyzed to support evaluation and decision making at all levels within the company. Thus, a company's performance and evaluation indicators need to focus on key results (Robin & Kaplan, 1991; Struebing, 1996; English, 1996).

Practices. Include those activities that occur within the organization as a means to achieve the strategic objectives.

Tools. Include different graphical and statistical methods for planning, collecting, analyzing, monitoring, and solving quality problems.

Some specific tools and techniques used in the TQM model may be different under each management practice. Some of the most commonly used tools and techniques (Graessel and Zeidler, 1993; Dean & Evans, 1994; St. Lawrence & Stinnett, 1994; Tedesco, 1994) are briefly described as follows: matrix diagrams, “are “spreadsheets” that graphically display relationships between ideas, activities, or other dimensions in such a way as to provide logical connecting points between each
item. A matrix diagram is one of the most versatile tools in quality planning. An Interrelationship Diagraph, "identifies and explores casual relationships among related concepts or ideas. It shows that every idea can be logically linked with more than one other idea at a time, and allows for "lateral thinking" rather than "linear thinking." A Tree Diagram, "maps out the paths and tasks necessary to complete a specific project or reach a specified goal. Thus, the planner uses this technique to seek answers to such questions as "what sequence of tasks will address the issue?" or "What factors contribute to the existence of the key problem?" A tree diagram brings the issues and problems revealed by the affinity diagram and the interrelationship diagraph down to the operational planning stage" (as cited in Evans & Lindsay, 1999, p.250-251). Quality Function Deployment (QFD), is an approach developed by the Japanese to meet customers' requirements throughout the design process and also in the design of production systems. According to (Graessel and Zeidler, 1993) "QFD is a customer-driven planning process to guide the design, manufacturing, and marketing of goods. Through QFD, every design, manufacturing, and control decision is made to meet the expressed needs of customers. It uses a type of matrix diagram to present data and information. Under QFD, all operations of a company are driven by the voice of the customer, rather than by edicts of top management or the opinions or desires of design engineers" (as cited in Evans & Lindsay, 1999, p. 405). Design of Experiments, is a test or series of tests that enables the experimenter to draw conclusions about the situation under study. It is used to improve the design of processes. For example, the Taguchi method is parameter design experiment aimed to reduce the variability caused by manufacturing variations. Taguchi categorizes
variables that affect the performance characteristics according to whether they are design parameters or sources of noise (as cited in Evans & Lindsay, 1999, p. 397-398).

The criteria used in the TQM model is the Malcolm Baldrige National Quality Award (MBNQA), which also includes the Criteria for Performance Excellence that establish a framework for integrating total quality principles and practices in any organization (as cited in Evans & Lindsay, 1999).

Another parallel between the BSC and TQM evaluation models, is that both of them employ a business performance scorecard. The TQM's performance scorecard includes a broad set of evaluation indicators that often consists of five key categories (Evans & Lindsay, 1999): First, customer satisfaction indicators (i.e., perceived value, overall satisfaction, complaints, gains and losses of customers, customer awards/recognitions). Second, financial and market indicators (i.e., return on equity, return on investment, operating profit, earnings per share, market share, percent of new product share). Third, human resource indicators (i.e., absenteeism, turnover, employee satisfaction, training effectiveness, grievances, suggestion rates). Fourth, supplier and partner performance indicators (i.e., quality, delivery, price, cost savings). Fifth, company specific indicators that support the strategy (i.e., defects and errors, productivity, cycle time, regulatory/legal compliance, new product introductions, community service, safety, environmental) (p.476).

TQM and Implementation Protocol

According to Ghobadian and Gallear (1997), the TQM's implementation process entails ten key steps (as cited in Hansson, 2003): First, recognition of the
need for the introduction of TQM. Second, development of an understanding among managers and supervisors. Third, establishment of goals and objectives of the quality improvement process. Fourth, development of a plan for TQM implementation. Fifth, training of the workforce. Sixth, creation of a systematic procedure. Seventh, alignment of the organization and development of a teamwork approach. Eight, implementation of the TQM concepts. Ninth, monitoring the implementation of TQM concepts. Tenth, engagement in continuous improvement by reestablishing new goals and objectives of the quality improvement process (p.36)

Six Sigma

Model Overview

The six sigma (6σ) is a business-driven, multi-faceted model to process improvement, reduced costs, and increased profits. With a fundamental principle to improve customer satisfaction by reducing defects, its ultimate performance target is virtually defect-free processes and products. The six sigma model, consisting of the following implementation steps "Define - Measure - Analyze - Improve - Control," (DMAIC) is the roadmap to achieving the customer improvement goal. Within this improvement model, it is the responsibility of the improvement team to identify the process, the definition of defect, and the corresponding measurements (Pyzdek, 2003).

The six sigma model originated at Motorola in the early 1980s in response to a CEO-driven challenge to achieve tenfold reduction in product-failure levels in five years. Meeting this challenge required swift and accurate root-cause analysis and correction. In the mid-1990s, Motorola divulged the details of their quality

57
improvement model, which since then has been adopted by several large manufacturing companies.

**Characteristics**

Conceptually, the sigma level of a process or product is where its customer-driven specifications intersect with its distribution. A centered six sigma process has a normal distribution with a mean, target and specifications placed six standard deviations to either side of the mean. At this point, the portions of the distribution that are beyond the specifications contain 0.002 ppm of the data (0.001 on each side). Practice has shown that most manufacturing processes experience a shift (due to drift over time) of 1.5 standard deviations so that the mean no longer equals target. When this happens in a six-sigma process, a larger portion of the distribution now extends beyond the specification limits (3.4 ppm).

The tools used in the six sigma evaluation model are often applied within a simple implementation process known as DMAIC. The DMAIC process is used when a project's goal can be accomplished by improving an existing product, process, or service.

As stated previously, the primary goal of six sigma is to improve customer satisfaction, and thereby profitability, by reducing and eliminating defects. Defects may be related to any aspect of customer satisfaction: high product quality, schedule adherence, cost minimization. Underlying this goal is the Taguchi Loss Function (Hurley & Loew, 1996), which shows that increasing defects leads to increased customer dissatisfaction and financial loss. Common six sigma metrics include defect rate (parts per million or ppm), sigma level, process capability indices, defects per

58
unit, and yield. Many six sigma evaluation indicators can be mathematically related to
the others (Pyzdek, 2003).

The six sigma evaluation model drive for defect reduction, process
improvement and customer satisfaction, and has the following characteristics:
everything is a process, all processes have inherent variability, data is used to
understand the variability and drive process improvement decisions.

**Evaluation Components and Evaluation Indicators used in Six Sigma Model**

Corresponding to the letters in the acronym DMAIC, this model’s five core
evaluation components are define, measure, analyze, improve, and control (Pyzdek,
2003):

**Define**. Under this component managers should define the goals of the
improvement activity. These goals are define not only by assessing customers’ needs,
but also from obtaining feedback from shareholders and employees. Goals include the
corporate, operational, and process level strategic objectives. Some of the underlying
questions included in this component are as follows: What is the business case for the
project? Who is/are the customer(s)? What is the current state map? What is the
future state map? What is the scope of this project? What are the deliverables? When
is the due date?

**Measure**. Under this component managers should measure the existing
system, by defining relevant and reliable evaluation indicators to help monitor
progress towards the previously defined goals. Some of the underlying questions
included in this component are as follows: What are the key metrics for this business
process? Are metrics valid and reliable? Do we have adequate data on this process?
How will the project leader measure progress?

**Analyze.** Under this component managers should examine the system or process to be improved in order to identify ways to eliminate the gap between the current and the desired performance. The analyses process starts by determining the current performance baseline of the system or process, and then descriptive data analysis is used to help managers understand the data. In addition, statistical tools are used to guide the analysis. Some of the underlying questions included in this component are as follows: What is the current state analysis? Is the current state as good as the process can do? Who will help make the changes? What are the resource requirements? What could cause this change effort to fail? What major obstacles does the project leader face in completing this project?

**Improve.** Under this component managers should improve the system or process by finding new ways to do things better, cheaper, or faster. Managers may use planning and management tools to implement the new approach, and also statistical methods to validate the improvement. Some of the underlying questions included in this component are as follows: What is the work breakdown structure? What specific activities are necessary to meet the project’s goals? How do the project leader will re-integrate the various subprojects?

**Control.** Under this component managers should control the system by institutionalizing the new process and aligning compensation and incentive systems, policies, procedures, budgets, operating instructions, and other management systems to the corporate strategic objectives. Managers may utilize standardization such as ISO 9000 to assure that documentation is correct. Additionally, managers may also
use statistical tools to monitor stability of the new systems or processes. Some of the underlying questions under this component are as follows: How do the project leader will control risk, quality, cost, schedule, scope, and changes to the plan? What types of progress reports should the project leader create? How do the project leader will assure that the business goals of the project were accomplished? How do the project leader will sustain the performance?

A similarity between the BSC and six sigma evaluation models, is that evaluation indicators of the six sigma model are based on the idea of a balanced scorecard. Balanced scorecards provide the means of assuring that six sigma projects are addressing key business results. Senior management are responsible for translating the stakeholders’ goals into evaluation indicators. These goals and evaluation indicators are then mapped to a strategy for achieving them. Scorecards are developed to display the evaluation indicators under each perspective. Finally, six sigma is used to either close gaps in critical indicators, or to help develop new processes, products, and services consistent with top management’s strategy (Pyzdek, 2003, p. 33-34).

For instance, if the goal in a six sigma project is to cut the time required to introduce a new product from 9 months to 3 months, some of the metrics that may be used include the average time to introduce a new product for the most recent month or quarter, and the number of new products introduced in the most recent quarter.

Some of the evaluation indicators used in the six sigma model under the four different BSC perspectives, are as follows (Pyzdek, 2003): First, financial indicators (i.e., cost per unit, asset utilization, revenue from new sources, profit per customer).
Second, customer satisfaction indicators (i.e., price, time, quality, selection, service relationship). Third, internal process indicators (i.e., product introductions revenue, key customer variables, inventory delivery costs, audit results for regulatory compliance). Fourth, learning and growth indicators (i.e., skills gaps from employee competencies, research deployment time from technology, and employee feedback from corporate culture). (p. 36).

Once an effort or project is defined, the six sigma team methodically proceeds through measurement, analysis, improvement, and control steps. The team is responsible for identifying relevant evaluation indicators based on engineering principles and models. Once the team has collected the data, then they may continue to analyze the data looking for trends, patterns, causal relationships and root causes for poor performance. Special experiments and modeling may be done in some cases to confirm hypothesized relationships or to understand the extent of leverage of factors; but many improvement projects may be accomplished with statistical and non-statistical tools. When the target level of performance is achieved, control measures are then established to sustain performance.

A partial list of specific tools to support each of the six sigma evaluation components are as follows: First, tools included in the define component are benchmark, baseline, kano model, voice of the customer, voice of the business, quality function deployment, process flow map, project management, and management by fact. Second, tools included in the measure component are defect metrics, data collection forms, plan, logistics, sampling techniques. Third, tools included in the analyze component are cause and effect diagrams, failure modes and
effects analysis, decision and risk analysis, statistical inference, control charts, capability, reliability analysis, root cause analysis, systems thinking. Fourth, tools included in the improve component are design of experiments modeling, robust design. Finally, tools included in the control component are statistical controls (i.e., control charts, time series methods), and non-statistical controls (i.e., procedural adherence, performance management, preventive activities). Additionally, process maps are created to show the linkage between suppliers, inputs, process activities, outputs, and customers. This technique is known in the six sigma model as SIPOC. The SIPOC technique helps identify those processes that have the greatest impact on customer satisfaction. Process maps are tools used in six sigma to provide managers with a picture of how work flows through the company. (Pyzdek, 2003, p. 67).

Six Sigma and Implementation Protocol

According to Pyzdek (2003), the steps required to successfully implement the six sigma model are as follows: First, educating senior managers on the philosophy, principles, and tools used in the six sigma evaluation model is critical. Additionally, managers should work on aligning and reducing the different organizational levels. Second, managers should develop systems to improve communication with customers, employees, and suppliers. This includes developing rigorous methods of obtaining and evaluating customer, employee, and supplier input. Third, managers should evaluate the skills of their employee teams, and then provide them with training on the philosophy, systems improvement tools, techniques used in six sigma. Fourth, managers should develop a model for continuous process improvement, along with a system of evaluation indicators for monitoring progress and outcomes. Six
Sigma metrics should focus on the organization's strategic goals, drivers, and key business processes. Fifth, managers should identify those business processes that need to be improved in the organization, aided by their employee teams and other people who have an adequate knowledge of these business processes. Six sigma projects should be conducted to improve business performance linked to measurable financial results. Finally, employee teams implement the different six sigma projects supervised by green belts and assisted by black belts project leaders.

Moreover, Gupta (2004) explains, "the current six sigma model consists of two implementation levels: the corporate level and the project level. Corporate level implementation requires leadership to take initiative and middle management to assist in developing a business case for adapting the six sigma model. The critical aspects of the corporate-level preparation for the six sigma model include establishing key business performance measurements, ensuring organizational effectiveness, assessing the organization's readiness for six sigma, and establishing goals for improvement. The project-level implementation relies on the Define - Measure - Analyze - Improve - Control (DMAIC) methodology to capitalize on opportunities for improvement" (p.37).

An important consideration throughout all the six sigma steps is to distinguish which process substeps significantly contribute to the end result. The defect rate of the process, service or final product is likely to be more sensitive to some factors than others. The analysis phase of six sigma can help identify the extent of improvement needed in each substep in order to achieve the target in the final product. It is important to note that six sigma performance (in terms of the ppm metric) is not
required for every aspect of every process, product and service. It is required only where it quantitatively drives a significant "factor" for the end result of customer satisfaction and profitability.

Institutionalizing the six sigma into the corporate culture might require significant investment in training and infrastructure. There are typically three different levels of expertise cited by Pyzdek (2003): "green belt, black belt practitioner, and master black belt. Each level has increasingly greater mastery of the skill set. Roles and responsibilities also grow from each level to the next, with black belt practitioners often in team/project leadership roles and master black belts often in mentoring/teaching roles." (p. 37).

Anatomy of Performance (AOP)

Model Overview

The AOP was developed by Geary Rummler in 2001. "AOP is a model that underlies an analytical approach that reflects the notion that organizations behave as systems. The AOP model identifies the major variables impacting individual performance and organization results, and it is based on three principles. First, every organization is a processing and adaptive system. The organization must be aligned. Second, every performer in an organization is in a human performance system. The human performance systems must be aligned. Third, the management system is key to keeping the performance system aligned. Management must be doing the aligning” (Rummler, 2002, p. 14).

“In order to diagnose where the “AOP” of a given organization is “broken” or misaligned, leading to sub-par performance, this situation is examined from four
views: management, business, performer, and organization system view. From this examination, the root causes of the poor performance in an organization are diagnosed in order to improve and sustain the desired performance” (Rummler, 2002, p. 14).

Some of the core components of the AOP model such as the “Human Performance System” (HPS), were first articulated by Rummler in 1964, while at the University of Michigan. It is a combination of B.F. Skinner’s work in reinforcement theory and basic industrial engineering practices. The development of the model was heavily influenced by Dale M. Brethower and George L. Geis, colleagues at the University of Michigan. The HPS is distinguishable from other “cause analysis” models because it conceptually and graphically recognizes the critical underlying principle that the variables impacting human behavior/performance are part of a system.

The AOP model has been employed in different organizations in successful performance, process, and evaluation improvement initiatives. Applications have spanned various service areas, including banking/financial, airline, automotive, telecommunications, hospitality, insurance, manufacturing, healthcare, pharmaceutical.

**Characteristics**

“The AOP model identifies the major variables impacting individual performance and organization results, and it is based on three principles (Figure 5): Under the first AOP principle every organization is a processing and adaptive system, and the organization system must be aligned.
Under the second AOP principle every performer in an organization is in a human
performance system, and the human performance systems must be aligned.

Under the third AOP principle, the management system is key to keeping the
performance system aligned, and managers must be doing the aligning” (Rummler,
2001, p. 17).

Some of the key points of the AOP model (Rummler, 2004), include: First,
organizations are systems, and every organization is a system that exists to produce
two types of system outputs (a) Desired products or services for some “receiving
system” or customer (b) An economic return to shareholders.

Second, organizations are processing systems: every organization is a
processing system of primary and support processes. Primary processes, are those
through which an organization produces a valued product or service (i.e., inventing,
developing, selling, and delivering products or services that directly impact the
customer). Support processes, are those that buttress the primary processes (i.e.,
human resources, finance, information technology).

Third, organizations are adaptive systems: an organization exists within a
larger system known as the ‘super-system.’ Elements of a ‘super-system’ include: the
consumer and capital markets, the resources/supply chain, the competition, and the
general business environment. Besides, every organization must adapt or die. The
organization must be able to accommodate changes in the larger super-system in
which it operates.

Fourth, jobs or roles and functions exist to support the processes of the
organization: all the tasks that make up the primary and support processes in an
organization are performed by a combination of individuals, machines, and computers. The tasks performed by individuals are usually organized into jobs, roles, or positions, which make up functions or departments. Functions and jobs should be linked to primary processes that add value to customers.

Fifth, all performers are part of a human performance system (HPS): each individual performer in any organization is also part of a unique personal system called the "human performance system." Components of a HPS are as follows: performer, input, output, consequences, and feedback. All components of the HPS must be in place at some minimal level if an organization is to get the desired performance from an individual in a consistent basis.

Sixth, management must keep the organization system aligned: management is essential to an organization adapting to its super-system, and keeping its internal processing system meeting customer expectations and organization goals. The failure of an organization to be aligned at any point in the AOP model is a failure of management. Effective management has three elements: First, the management system or infrastructure is made up of processes and procedures. Second, the management skills, as exemplified by the ability to work effectively within the management system to deliver desired results. Third, includes leadership that consists primarily of setting appropriate direction and enrolling the organization in following that direction.

Seventh, the results chain must link to a critical business issue within the AOP model of any organization, a usually invisible results chain links these three primary levels of performance or results: (a) Organization-level performance or results,
related to expectations of stakeholders and customers (the two primary receivers of organization outputs), (b) Process-level performance or results, which are necessary for the organization to produce its outputs and meet the expectations of customers and stakeholders, (c) Job-level performance or results, which are necessary for primary and support processes to achieve their goals.

Figure 5. Anatomy of Performance (AOP) model

The Anatomy of Performance is a scaleable model that applies to the: (a) total company, (b) division or business unit, (c) plant or district, (d) department. Additionally, the AOP include the following attributes: First, customer needs are
aligned with shareholders needs. Second, organization goals are aligned with the reality of its super-system (or larger “business system”). Third, primary processes are aligned to meet customer expectations and organization goals. Primary processes are those having to do with inventing, developing, selling and/or delivering products/services and directly impact the customer. Fourth, support processes are aligned with Primary process goals. Support processes are those that support the primary processes and are typically related to Human Resources, Finance, Information Technology. Fifth, function/jobs/roles are aligned to perform the required tasks of the processes. Sixth, the human performance system components are aligned (individually, vertically, and horizontally). Seventh, management is doing the aligning.

“Performance analysis is about overlaying the “should” Anatomy of Performance template on an organization’s “is” reality, identifying differences from the “should” elements, assessing the likely impact of the differences on the target gap in results, and specifying changes to close the gap in results” (Rummler, 2003, p.64).

**Evaluation Components and Evaluation Indicators used in AOP Model**

Moreover, Rummler (2003) explains that in order to diagnose where the “Anatomy of Performance” of a given organization is “broken” or misaligned, leading to sub-par performance, this situation is examined from four views (Figure 6):

- **Management View.** This view addresses performance principle number three in which management must do the aligning. The goal of this view is to assess the quality of the management being provided to the client entity. Under this view it is
important to assess three dimensions of management including infrastructure or
system, culture, and quality of leadership as set and executed by senior managers.

**Business View.** There are two aspects of the business view. The first is basic
background about the company, its industry, ownership and performance history.
These are facts that can not be changed, but usually provide some insight into
possible constraints on what “could be” in the future. Second are those things that
reflect important business decisions or assumptions made by company management,
such as direction, key performance variables, economic model and business values.

**Performer View.** This view pertains to performance principle number two,
which states that “human performance systems must be aligned.” Under this view it
is important to identify performers who are critical to successfully closing the gap in
results but whose “is” behavior and/or performance will need to be changed in order
for them to do so. Thus, it is important to specify the “should” behavior or
performance for those individuals, determine the factors that support the “is” state
and specify the changes necessary to get and sustain the “should” behavior and/or
performance.

**Organization System View.** This view addresses performance principle
number one, which states that “the organization system must be aligned.” The
organization system must be aligned from the super-system down to the individual
performer. “A super-system is the larger system in which our target system exists. If
the system in question is the company, then its super-system consists of the
product/service market, the shareholders, competition, resources, and general
business environment” (Rummler, 2003., p. 8). In this view, it is important to
examine the “is” state of each level in this system, as well as the alignment between the levels.

Figure 6. Four views of an organization

Many types of data and information are needed for the AOP performance improvement model, including: customer needs, product and service performance, operations performance, market assessments, competitive comparisons, supplier performance, employee performance, and cost and financial performance. A major
consideration for the AOP model is to identify and manage those key performance variables or indicators that impact the success of an organization, and to answer the question of how the variables impact the critical business issues or results gap.

The AOP model, collects and analyzes data that includes either four data sweeps for large-scale organization analysis projects, or two to three sweeps for smaller projects. The steps during this data sweep(s) collection are as follows: (1) data sweep planned, (2) client update and discussion, (3) data gathered, (4) data analyzed, (5) changes specified.

The evaluation indicators used should best represent the factors that lead to improved customer, operational, and financial performance. Thus, a company’s performance measurements need to focus on key results. Some specific tools and techniques used in the AOP model may be different under each phase of the results improvement implementation process, but some of the most commonly used tools and techniques are briefly described as follows: performance logic map, this is a useful tool for assessing the impact that components of a process have on desired results. Human performance system analysis and improvement guide, this is a valuable tool for determining the causes of poor performance and developing a plan to correct them. Initiative Analysis and Management, this tool is used for evaluating and managing potentially competing initiatives. Value chain impact matrix, this is a useful tool for assessing the impact that components of the value chain have on desired results. Cross-Functional Process Map, this tool is used to document how a process “cuts-across” organization functions. Organization level performance logic
map, this is a helpful tool for determining what performance variables are impacting the gap in results.

**AOP and Implementation Protocol**

The results improvement implementation process has four phases:

First, desired results should be determined and the project defined. Objectives under this phase are as follows, (a) to determine if there is a significant results gap to be closed, (b) to determine the feasibility of closing the results gap, (c) to prepare a project plan for closing the results gap. Some of the questions included in this first phase are: What and where is the gap in results? Is the gap significant? Is it feasible to close the gap?

Second, barriers should be determined and the changes specified. Objectives under this phase are as follows: (a) to determine what components of the four views must be realigned to close the results gap, (b) to specify the changes required in the four view components to close the results gap. One question included in this second phase is: Why the gap in results and what is required to close it?

Third, changes should be designed, developed, and implemented. An objective under this phase is to design, develop, and implement the interventions necessary to close the results gap and assure continuous improvement. One question included in this third phase is as follows: How are we closing the gap in results?

Fourth, results should be evaluated and maintained or improved. One objective under this phase is to determine if the results gap has been closed and if not, what must be done to do so. One question included in this phase is as follows: Did we close the gap in results?
A detailed summary of the AOP results improvement implementation process is provided in Table 4.

Table 4.

Summary of the AOP results improvement implementation process

<table>
<thead>
<tr>
<th>PHASE I</th>
<th>PHASE II</th>
<th>PHASE III</th>
<th>PHASE IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desired results determined and project defined</td>
<td>Barriers determined and changes specified</td>
<td>Changes designed, developed, and implemented</td>
<td>Results evaluated and maintained or improved</td>
</tr>
<tr>
<td>Objective:</td>
<td>Objective:</td>
<td>Objective:</td>
<td>Objective:</td>
</tr>
<tr>
<td>• Determine if there is a significant Results Gap to be closed.</td>
<td>• Determine what components of the Four Views must be realigned to close the Results Gap.</td>
<td>• Design, develop and implement the interventions necessary to close the Results Gap and assure continuous improvement.</td>
<td>• Determine if the Results Gap has been closed and if not, what must be done to do so.</td>
</tr>
<tr>
<td>• Determine the feasibility of closing the Results Gap.</td>
<td>• Specify the changes required in the Four View components to close the Results Gap.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Prepare a Project Plan and proposal for closing the Results Gap.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outputs:</td>
<td>Outputs:</td>
<td>Outputs:</td>
<td>Outputs:</td>
</tr>
<tr>
<td>• Project Definition Worksheet (PDW)</td>
<td>• Findings and Recommendations Worksheet</td>
<td>Implemented Changes</td>
<td>Continuously Improved Performance</td>
</tr>
<tr>
<td>• Project Plan</td>
<td>• Macro Design and Implementation Plan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Proposal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions To Answer:</td>
<td>Questions To Answer:</td>
<td>Questions To Answer:</td>
<td>Questions To Answer:</td>
</tr>
<tr>
<td>1. What and where is the Gap in Results?</td>
<td>Why the Gap in results and what is required to close it?</td>
<td>How are we closing the Gap in Results?</td>
<td>Did we close the Gap in Results?</td>
</tr>
<tr>
<td>2. Is the Gap significant?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Is it feasible to close the Gap?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major Steps:</td>
<td>Major Steps:</td>
<td>Major Steps:</td>
<td>Major Steps:</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>2. Results Gap determined</td>
<td>2. Data Sweeps executed and results analyzed</td>
<td>2. Changes designed and developed</td>
<td>2. Deviations from expectations analyzed</td>
</tr>
<tr>
<td>3. Feasibility assessed</td>
<td>3. Client apprised of progress and issues</td>
<td>3. Changes “pilot tested” when appropriate</td>
<td>3. Modifications made as necessary:</td>
</tr>
<tr>
<td></td>
<td>5. Recommendations summarized</td>
<td>5. Changes installed, monitored and supported</td>
<td>b. To implementation of initial changes</td>
</tr>
<tr>
<td></td>
<td>6. Macro Implementation Plan developed</td>
<td></td>
<td>4. Conclusions reached regarding effectiveness of initial solutions</td>
</tr>
<tr>
<td></td>
<td>7. Prototypes developed as appropriate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tools:</th>
<th>Tools:</th>
<th>Tools:</th>
<th>Tools:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Problem Pentagon</td>
<td>(17) Process Impact Matrix</td>
<td>(36) Basic Change Model for Closing Gap in Results</td>
<td></td>
</tr>
<tr>
<td>(2) Super-System Map</td>
<td>(18) Process Performance Table</td>
<td>(37) Understanding Impact of Recommendations</td>
<td></td>
</tr>
<tr>
<td>(3) Super-Duper System Map</td>
<td>(19) Components of an Effective Work Process</td>
<td>(38) Rating Past Implementation Efforts</td>
<td></td>
</tr>
<tr>
<td>(5) “IT” Business Organization Model</td>
<td>(22) PMMS and Organization Hierarchy</td>
<td>(39) Implementation Step Model</td>
<td></td>
</tr>
<tr>
<td>(6) Value Chain &amp; Function View</td>
<td>(24) Linking Processes, Functions, and Jobs</td>
<td>(40) Initiative Analysis and Management</td>
<td></td>
</tr>
<tr>
<td>(7) Cross-Functional Value Chain Map</td>
<td>(25) Job Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) Value Chain Impact Matrix</td>
<td>(26) Troubleshooting the HPS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10) Linear Process Map</td>
<td>(27) HPS Template</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11) Cross-Functional process Map</td>
<td>(28) HPS Guide and Worksheet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(13) Performance Logic Map</td>
<td>(29) HPS Alignment Templates</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summary and Comparison of Evaluation Models

The BSC, CIPP, TQM, Six Sigma, and AOP models are often quoted as being alternative ways of informing and improving both strategic and operational management decision-making. These evaluation models differ in their orientation, the types of variables, information requirements, implementation protocols, and outcomes. However, all of these different evaluation models have a common purpose. They are all used to implement strategic performance evaluation that facilitate managers’ strategic decision-making, planning and control.

Moreover, when analyzing these evaluation models, it is also important to note that these different models seem to be very similar regarding aspirations, and concepts. Indeed, one can probably agree that these approaches share a number of characteristics. They are all measurement based, they encourage a dialogue about strategic decision-making and performance improvement, they all strive to act as catalysts for change and action, and all are based on principles of on-going review, learning, and feedback. Above all, long-term success in implementing one or a combination of these different models depends on management’s on-going commitment to improve an organization’s performance.

The BSC is consistent with the CIPP model as both are decision/oriented approaches intended to provide information to people in organizations to facilitate managers’ strategic decision-making, planning, and control. The BSC model builds on key concepts of evaluation practice that can be found in the CIPP model, including customer-defined (i.e., meeting stakeholders needs), continuous improvement, emphasis on organizational effectiveness, and are measurement-oriented management.
models. For instance, efforts to improve the quality, responsiveness, and efficiency of internal processes that can be found in the process evaluation core part of the CIPP model can be reflected in the operations portion of the BSC's internal perspective. Thus, companies already implementing different evaluation models in their initiatives will find ample opportunity to sustain their programs within the more strategic BSC or CIPP models.

The intended uses of BSC and CIPP models are similar. Both the BSC and CIPP focus primarily on improvement, accountability, and enlightenment. Improvement involves providing information for assuring the quality of a service or for improving it (Kaplan & Norton, 1996, p. 31). Close attention is given to the needs of stakeholders, and to the link between process and outcome. The second main role of both the BSC and CIPP models is to produce accountability or summative reports. The BSC is a valuable tool for accountability purposes, and for broadening relationships with stakeholders (Kaplan & Norton, 1996, p. 31). The CIPP model serves not only to guide operating programs and to summarize their contributions, but also to improve them (Stufflebeam, Madaus, & Kellaghan, 2000, p. 173) "We cannot be sure that our goals are worthy unless we can match them to the needs of the people they are intended to serve" (Stufflebeam, Madaus, & Kellaghan, 2000, p. 188). The CIPP model not only fosters improvement, but also provides accountability records. The third use is enlightenment, in which the BSC and the CIPP models attempt to consider all criteria that apply in for determination of value.

The BSC is consistent also with TQM principles. Initiatives to improve the quality, responsiveness, and efficiency of internal processes that can be reflected in
the operations portion of the BSC's internal perspective. Extending TQM principles out to the innovation process and to enhancing customer relationships will be reflected in the several other building blocks in the internal business process perspective. Thus, companies already implementing the continuous improvement and measurement disciplines from TQM will find ample opportunity to sustain their programs within the more strategic framework of the BSC.

However, the BSC does much more than merely reframe TQM principles into a new model. The BSC enhances, in several ways, the effectiveness of TQM programs. First, the BSC identifies those internal processes where improvement will be most critical for strategic success. In many organizations, TQM programs succeeded; but, their impact could not be detected in the financial or customer performance of the organization. The BSC identifies and sets priorities on which processes are most critical to the strategy, and also focuses on whether the process improvements should center more on cost reduction, quality improvement, or cycle-time compression.

Another difference from the BSC to TQM programs occurs by forcing managers to explicate the linkage from improved operating processes to successful outcomes for customers and shareholders. Companies focusing only on quality and local process improvement often do not link operational improvements to expected outcomes in either the customer or the financial perspective. The BSC requires that linkages be explicit. One linkage is from quality improvements in the internal perspective to one or more outcome (not process) measures in the customer perspective. A second link is from quality improvements that enable companies to
reduce costs, an outcome in the financial perspective. The BSC model enables managers to articulate how they will translate quality improvements into higher revenues, fewer assets, less people, and lower spending.

It is important to mention that there is no real difference between six sigma and the TQM models. Indeed, six sigma does employ some of the same tools and techniques the TQM models. Indeed, six sigma does employ some of the same tools and techniques of TQM. Both six sigma and TQM emphasize the importance of top-down support and leadership. Both models make it clear that continuous quality improvement is critical to long-term business success. In addition, the plan-do-study-act cycle used in TQM is not fundamentally different than the Six Sigma’s define-measure-analyze-improve-control model’s core parts. However, there are also some differences between these two models, such as: six sigma extends the use of the improvement tools to cost, cycle time and other business issues. Six sigma integrates the goals of the organization as a whole into the improvement effort. Certainly, quality is good, but not independent of other business goals. Six sigma creates top-level oversight to assure that the interests of the entire organization are considered.

Concerning, the AOP and BSC models, it is possible to say that AOP is in agreement with the BSC model regarding the need for managers to have a set of instrument panels to review. Kaplan and Norton call it “balance.” On the other hand, Rummler calls it tracking the variables that impact the performance of a business system. These are the following instrument panels suggested by Rummler:: (a) “tracking external variables, as represented by the external components of the super-system, (b) tracking the financials, (c) tracking critical success factors and/or
operating factors (e.g., market share) as determined by the strategy, (d) tracking critical resource utilization, such as human resources, technology. However, the specific instrument panels and meters in those panels will vary with the organization, based on its strategy and particular industry position” (Rummler, 2002, p. 233).

Moreover, Rummler (2002) observes that “the meters in instrument panels b) through d) above, must be linked to provide an integrated set of measures or instrument panels. Rummler explains that there is an underlying logic or set of logics which link most aspects of an organization’s performance, and that the major point that all these meters are linked is at the processes, or at the Process Level” (Rummler, 2002, p. 233).

Furthermore, Rummler adds that there is a “performance logic” inherent in every process, and that it is also an Organization Level “performance logic” (one or more) which links the various processes in an organization.

In order to link the three levels of performance (Organization, Process, and Job/Performer Levels), and to get consistent, high performance in an organization there needs to be an underlying logic or Performance Logic. According to Rummler and Brethower, “A Performance Logic (PL), is a network of variables or factors that affect a given output. In addition, not all variables in the Performance Logic are “born equal.” Some are more critical to the desired output than others. These variables are called “Leverage Points”- those variables in the PL which will have the greatest impact on the desired output and should therefore be measured, monitored, and managed. However, any given company in an industry may select particular Leverage Points that they will emphasize in order to give them a competitive edge”
In addition, “the performance logic is what gives managers information on: First, what performance is required at all levels of the logic. Second, what performance to monitor. Third, what performance to measure. Fourth, what questions to ask about performance deviations. Fifth, what actions to take to modify performance” (Rummler, 2001, p.121).

Furthermore, Rummler (2002) explains that “the enterprise measures (those measures by which the senior executives of the organization choose to measure the success of the enterprise) are the starting point of the design of a performance measurement system. These enterprise measures should be influenced by the strategy. Once the enterprise measures are known, then it is possible to develop the ‘performance logics’ that link the enterprise measures to the core processes. As measures are developed for the critical leverage points in the performance logic, then these measures are linked to the processes and to each other” (p.233).

As a final point, Rummler (2002) adds that “performance measures that does not take into account the performance management system to be employed, is limited. It is not possible “to develop a good measurement system (which provides data) without a performance management system (which specifies what information is needed from the data to assist in making management decisions). The management system provides the management decision-making context for the measurement system. A measurement system without a performance management system is incomplete” (p. 234).
CHAPTER III
METHODOLOGY

The purpose of this chapter is to detail the methodology used for answering the three remaining research questions (2-4) posed by this study:

2. What are the similarities and differences related to actual implementation of BSC and CIPP Models in terms of its methods, including: evaluation components, evaluation indicators, data collected to support the evaluation, evaluation implementation protocol, and qualitative and quantitative analyses?

3. What are the outcomes of these two (BSC and CIPP) evaluation models; what are the similarities and differences?

4. What are the critical factors that seem to be associated with successful applications of the BSC and CIPP Model?

A detailed description of the study’s research strategy, including the use of multiple case study research method (Yin, 1994), is presented and justified. Second, the Success Case Method (Brinkerhoff, 2003), is described as the evaluation tool used to examine the different case studies in organizations that have implemented the BSC and CIPP evaluation models. Included in this discussion, are statements of the three remaining research questions proposed in this study. Third, a detailed description of the selection and description of the BSC and CIPP’s case studies used in this study is provided. The section concludes with a description of the steps used to analyze each of these case studies.
Case Study Research Methodology

This study was based on a multiple case study research method describing the context/applications, uses, methods, products, of the BSC and CIPP models, supported by the experiences of those evaluators and practitioners in organizations that have implemented them. The use of case study research and the Success Case Method (SCM) as an evaluation tool, were seen as being the most tenable strategies to answer the research questions 2, 3, and 4 posed by this study. The case studies selected in this study were used as a means to describe, understand, and examine the similarities and differences of those organizations that have implemented the BSC and CIPP models.

Some of the formal definitions of case studies that were found in the literature, include the following: The United States General Accounting Office (GAO) Program Evaluation and Methodology Division (1990) defines case studies as “A case study is a method for learning about a complex instance, based on a comprehensive understanding of that instance obtained by extensive description and analysis of that instance taken as a whole and in its context” (p. 14). According to Yin (1994) “In general, case studies are the preferred strategy when “how” and “why” questions are being posed, when the investigator has little control over events, and when focus is on a contemporary phenomenon in some real-life context” (p.1). Moreover Schramm (1971) “the essence of a case study, the central tendency among all types of case study, is that it tries to illuminate a decision or set of decisions: why they were taken, how they were implemented, and with what result” (as cited in Yin, 1994, p. 12).
Although case studies are valued by a great number of researchers, there is a good deal of variability in their uses. For example, In the Handbook of Qualitative Research (Denzin and Lincoln, 1994), contributing author Robert Stake identified three types of studies inherent in case study research: (1) intrinsic case study, undertaken to gain better understanding of a particular case; (2) instrumental case study, undertaken to provide insight into a particular issue or for the refinement of theory; and (3) collective case study, whereby an instrumental study is extended to several cases (pp 237-238). The GAO in their publication Case Study Evaluations (1990), identified the following six types of case studies: (1) Illustrative, which is a descriptive case study that adds in-depth examples to other information about a program or policy; (2) exploratory, which is descriptive but aims to generate hypotheses for later investigation; (3) critical instance, which examines a single instance of unique interest or serves as a critical test of an assertion about a program, problem or strategy; (4) program implementation, which is usually a normative investigation of operations at several sites; (5) program effects, which examines causality and multisite, multimethod assessments; and (6) cumulative, which brings together findings from many case studies to answer evaluation questions be they descriptive, normative or cause-and-effect.

All of these depictions of case studies define the study at hand in various ways. In agreement with Stake’s utility classification of case study research, this is an intrinsic case study type, given that the information obtained from each of the BSC and CIPP’s case studies that will be examined in this study will be used to gain a better understanding of some of the uses, methods, outcomes and critical factors of
those organizations that have implemented these evaluation models. Additionally, this was a cumulative case study research type (GAO, 1990) as it aimed to bring together findings from different case studies to answer the three remaining research questions posed by this study. The current study was attempting to determine how the use of the BSC and CIPP evaluation models when used in organizations as strategic management systems, can benefit profit and nonprofit organizations by achieving long-term strategic objectives, implementing strategies and linking them to unit and individual goals. The current study also sought to gain a comprehensive understanding of the evaluation tools and methods that are used in both models and can be integrated and applied selectively in performance evaluation contexts are used in both models. This study looked also to provide guidance to evaluators by devising better alternatives and solutions to reach the desired program's outcomes that can be obtained from the BSC and CIPP evaluation models. Finally, this study aimed to gain an understanding of what were the strengths and weaknesses of each model to examine the critical factors that seem to be associated with successful applications of these evaluation models.

With this in mind, the multiple case study method was seen as an optimal strategy for investigating the original dynamics evident in the corporate use of the BSC and CIPP models as powerful tools for organizational evaluation by providing managers with information about "what performance is required at the organization, process, and job/performer level, what performance to measure, what questions to ask about performance deviations, and what actions to take to modify performance" (Rummler, 2001, p. 113). The use of multiple case study method was also viewed as
an ideal way for capturing the experiences of BSC and CIPP evaluators and practitioners, which could provide substantial data.

**Multiple Case Study Design**

Although there are disadvantages to the use of multiple case study designs, there are distinct advantages. One of which is the use of multiple cases designs which increases generalization of results by replicating the pattern-matching, thus increasing confidence in the robustness of the theory. Notably, the evidence from multiple cases is often considered more compelling, and the overall study is regarded as being more robust. A disadvantage often levied against case study methodology is that its dependence on a single case renders it incapable of providing a generalizing conclusion (Yin, 1994). In addition, Yin (1994) noted that another method to overcome the issue of generalization of results in case study designs, is by “the goal of the study should establish the parameters, and then should be applied to all research. In this way, even a single case could be considered acceptable, provided it met the established objective” (p. 46). Moreover, Yin (1994) pointed out that generalization of results, from either single or multiple designs, is made to theory and not to populations. Other disadvantages when using multiple case study designs is that they are usually more expensive and time-consuming to conduct than single-case designs.

This study used a multiple case study design, and qualitative data from these case studies was analyze to examine the context/applications, uses, methods, products, strengths, weaknesses, and limitations resulting from the implementation of the BSC and CIPP models in different organizations.
Success Case Method

The use of a “case study protocol” is essential when using a multiple-case design. Yin (1994) recommended the use of case-study protocol as part of a carefully designed research project. Through the review of published case studies, and also by employing the Success Case Method (SCM) (Brinkerhoff, 2003) as an evaluation method to conduct a multiple case study design, each of the case studies was analyzed in this particular study by finding out how well the BSC and CIPP models have worked in those organizations that have implemented them successfully.

“The Success Case Method is a useful evaluation tool to measure the impact of any performance improvement initiative. With this evaluation method, we can identify, document, and quantify specific instances of positive performance impact as a result of our learning solution. We can also identify environmental factors that can impede performance, helping us to diagnose issues of transfer, and work to improve our solutions. When view in this manner, evaluation becomes a vital tool to help us improve the value of our performance solutions, not a report card that validates the worth of the training department” (Brinkerhoff, 2003, p. 20).

The SCM guided the execution of the multiple case study design proposed in this study by identifying those BSC and CIPP “success cases” that organizations have implemented, and second by providing a framework to guide the analysis for each of these cases.

To conduct the study using the SCM, first “BSC and CIPP successful cases” were identified from those organizations that have implemented these evaluation models. Therefore, the use of the SCM helped to be selective, and to focus on a few
successful BSC and CIPP case studies that portrayed the critical key factors or issues that were fundamental to understanding the implementation process and outcomes of these evaluation models. Brinkerhoff (2003) found the following:

The Success Case Method likewise leverages the measurement approach of analyzing extreme groups, because these extremes are masked when mean and other central tendency measures are employed. This is the same concept applied in Shainan quality methods that are employed in some manufacturing operations to assess and diagnose quality of machine parts. The Shainan method directs quality assessors to analyze a sample of the very best parts produced by a manufacturing process as well as a sample of the very worst. From these extreme samples, manufacturing process elements are targeted and revised to assure a greater consistency of only “BOB” (the best of the best) parts and reduce the frequency of “WOW” (worst of the worst) parts. (p. 17)

Moreover, the Success Case Method encompass a two-part structure: First, a survey may be used to find potential and likely “success cases” including those individuals (or teams) that seem to be the most successful in using some new change or method. Another approach to identify potential success cases may be by reviewing usage records and reports, accessing performance data, or simply by asking people. Second, an interview may be conducted to identify and document the actual type of success being achieved. The interview starts by trying to determine whether the person interviewed truly represents a success case. Assuming that this is true, the interview then proceeds to probe, understand, and document the success. This interview
provides information about how those teams or individual obtained results from integrating the new change or method in their work. By conducting these interviews, evidence-based data should be collect to confirm those success cases found in the organization.

This study followed the five steps proposed by Brinkerhoff (2003) in order to use the SCM method as a guide for the execution of the multiple case study design:

1. **Focusing and planning a Success Case study.** Chapter I and III contains the information included in this first step such as: the background and purpose for conducting this study, the research questions that guided this work, the relevance of this study, and definitions. Chapter III includes information on: the research strategy that was used (multiple case study design), the method that was used (SCM evaluation method) to analyze the information, and the organizations (participants) that composed the sample for this study.

2. **Creating an “impact model” that defines what success should look like.**
   The impact model in this particular study is provided in Chapter II containing information on what are the intended uses and outcomes of the BSC and CIPP evaluation models if they are implemented well. The information from Chapter II was used to compare it with the identified BSC and CIPP success cases.

3. **Designing and implementing a survey to search for best and worst cases.**
   A survey was not use for this particular study. Case studies were identify by using the SCM method and reviewing literature on the topic of BSC
4. **Interviewing and documenting success cases.** This part of the SC study was accomplish by asking evaluators and practitioners to provide information on those documented success cases and not success cases in order to identify those factors that made success possible. Besides, Brinkerhoff (2003) noted that “Almost always, an SC study also looks at instances of nonsuccess. That is, just as there is some small extreme group who has been very successful in using a new approach or tool, there is a likewise some small extreme group at the other end who experienced no use or value. Investigation into the reasons for lack of success can be very enlightening and useful. Comparisons between groups are especially useful” (p.17).

Moreover, the SCM guided this study by providing a useful case study protocol for analyzing each of the case studies by focusing on these particular questions:

- What is really happening, and not happening, as a result of implementing the BSC and CIPP models in different organizations?
- What results are the BSC and CIPP models helping to produce? The review of the different case studies will help to search for evidence about the most moving and compelling results that BSC and CIPP are producing.
• What is the value of the results? To make decisions about how much more value the BSC and CIPP are realistically capable of making above and beyond its current level of impact.

• How could the BSC and CIPP models be improved? To assess those factors or key issues that are associated with success.

5. **Communicating findings, conclusions, and recommendations.** Chapter IV and V contained information related to the research questions posed in this study and a discussion of issues related to BSC and evaluation models’ practices, and presented recommendations for evaluators and researchers.

**Selection and Description of BSC and CIPP's Case Studies**

The case studies selected contained data from organizations that have implemented the BSC and CIPP models from 1995-2005.

Because one of the goals of this study was to provide guidance to those evaluators and practitioners interested in getting some direction about the different evaluation models, including an understanding of these distinctions in context/applications, uses, methods, products, strengths, weaknesses, and limitations. Case studies were used in this study to depict a holistic portrayal of those organizations that have implemented these evaluation models in order to learn from their experiences and results regarding the effectiveness of each of these models.

The data from these case studies examined some of the evaluation components, evaluation indicators, data collected to support the evaluations,
implementation protocols, qualitative and quantitative analyses, outcomes, and critical factors associated with successful implementations of the BSC and CIPP evaluation models.

Sample

The data used in this study was collected from different sources (Appendix A) and organizations from a range of industries including: Siemens AG in Germany (Power, Automation and Control, Transportation, Medical, Lightning and Information and Communications); Hilton Hotel Corporation (Hospitality/Services); Mobil North America marketing and refining (Oil corporation); United Parcel Services, UPS (Transportation Company); The Spirit of Consuelo: An Evaluation of Ke Aka Ho’Ona (Community Development and Self-Help House Construction); NASA-AESP (Aeronautics and Aerospace Industry).

Some of the above BSC’s case studies including: Siemens AG in Germany, Mobil NAM&R, Hilton Hotel Corporation, and United Parcel Services were found at “The Balanced Scorecard Collaborative.” All three of these case studies are part of a collection of ‘Hall of Fame Case Studies’ that have been published in the Balanced Scorecard Collaborative webpage. These case studies were selected because they constitute exemplary illustrations of those organizations that have successfully implemented the BSC methodology. According to the Balanced Scorecard Collaborative, “Members of the Balanced Scorecard Hall of Fame exemplify best-practice Balanced Scorecard (BSC) implementation. Members of the BSC Hall of Fame have used the BSC to become Strategy-Focused Organizations and to achieve breakthrough results. Hall of Fame members consist of organizations from a wide
variety of industries that are geographically dispersed throughout the world, and range in size from 200 employees to more than a million" (Balanced Scorecard Collaborative, 2004).

In addition, “Balanced Scorecard Collaborative Hall of Fame winners have achieved breakthrough performance largely as a result of applying one or more of the five principles of a Strategy-Focused Organization: mobilize change through executive leadership, translate the strategy to operational terms; align the organization to the strategy; make strategy everyone’s job; and make strategy a continual process. Other selection criteria are: implement the Balanced Scorecard as defined by the Kaplan/Norton methodology; present the case at a public conference; achieve media recognition for the scorecard implementation; produce significant financial or market share gains; and demonstrate measurable achievement of customer objectives” (Balanced Scorecard Collaborative, 2004).

These BSC’s case studies presented in this study were selected from a handful of case studies, because they represent good examples of practitioners that have used the BSC and other evaluation models including TQM and Six Sigma in different organizations. Besides, each of these case studies provides some lessons learned from their successes and failures while implementing the BSC model; thus providing ideas to practitioners on how they can improve their organization’s adoption of this model.

Case studies were selected in order to address specific issues concerning BSC implementation process and results (i.e., BSC and CIPP design issues, implementation process and principles, integration of different evaluation models such as BSC and Six sigma, use of elements, BSC software), link between BSC and
compensation. Issues concerning with merging BSC with other organizational solutions and different models were covered in Siemens AG and United Parcel Services (UPS) case studies.

The CIPP’s case studies presented in this study (The Spirit of Consuelo: An Evaluation of Ke Aka Ho’ona, 2002, and NASA-AESP, 2004) were found at the Evaluation Center Library source. These case studies were selected from a handful of case studies, because they constitute not only an exemplary illustration of those organizations that have successfully implemented the CIPP model, but also they represent good examples of practitioners that have used the CIPP evaluation model. Each of these case studies provide some lessons learned from their successes and failures while implementing the CIPP model; thus providing ideas to practitioners on how they can improve an organization’s adoption of these evaluation models.

Data Preparation

The data preparation approach used in this study involved four steps: First, an abstract looking at general information including basic background about the company, its industry, ownership, and performance history was included. Second, a statement of the problem that the organization was facing at the time the executive team decided to implement the BSC or CIPP evaluation models, and a brief explanation of the rationale to implement BSC or CIPP as a solution. Third, information regarding the BSC or CIPP implementation process was analyze under this step. Additionally, information on how each of the organizations developed the BSC or CIPP model and how these models were used was included. Moreover, the methodology used, and integration of BSC or CIPP with other methodologies or
process improvement mechanisms, and elements used in the construction and implementation of BSC or CIPP was reviewed and analyzed. Finally, information on outcomes, as well as lessons learned in each of the case studies was reviewed under this issue.

Analysis

The SCM was used as the analytic strategy for this study and followed the five SCM steps method described above to answer the research questions posed by this study. Specifically, the SCM method provided an organizational structure for understanding the differences and interrelationships between BSC and CIPP evaluation models, including the distinctions in terms of: methods, products, strengths, weaknesses, limitations, of each of these evaluation models.

The methodology used to analyze the content data from each BSC and CIPP case study included the following steps:

First, each individual case study was reviewed using the SCM case study protocol to guide the identification of information related to BSC and CIPP implementation process and results. The information was coded into a set of categories that were relevant to the research questions included in this study.

Second, data from each case study consistent with each of the categories identified was coded. Different coding categories were included, such as: types of evaluation components included in each model (i.e. evaluative information obtained in each component), type of evaluation indicators (qualitative, quantitative, or both; coverage of different organizational performance areas such as customer, operational, process, financial), type of data collected, implementation protocol (i.e., principles,
procedures, time requirements, challenges), qualitative and quantitative analyses (i.e.,
methods and type of tools), outcomes (i.e. intended and unintended, positive and
negative), and critical factors (i.e., alignment of evaluation indicators, integration of
different evaluation models, planning, link of evaluation models with other
management systems including compensation and appraisal system).

Third, categories were analyzed by using a pattern-matching logic. Yin
(1994) observed “For case study analysis, one of the most desirable strategies is to
use a pattern-matching logic. Such a logic (Trochim, 1989) compares an empirically
based pattern with a predicted one (or with several alternative predictions). If the
patterns coincide, the results can help a case study strengthen its internal validity”
(p.106).

Fourth, the overall pattern of results from each of the four BSC cases were
compared with one another. The same procedure was used to analyze the pattern of
results from the two CIPP cases included in this study.

Fifth, the findings from these analyses were included in a written report for
each individual BSC and CIPP case study.

Summary tables were used to facilitate the analysis and to emphasize some of
the context/applications, uses, methods, products, strengths, weaknesses in each of
the BSC and CIPP case studies reviewed.

Summary

The methodology used in this study was appropriate based on the problem
identified and the rationale for this study. The following summarizes the critical steps
and decisions used in this study.
1. Multiple Case Study Design. This study used a multiple case study design, and qualitative data from these case studies was analyzed to examine the evaluation components, evaluation indicators, data collected to support the evaluation, implementation protocol, qualitative and quantitative analyses, outcomes, and critical factors associated with successful implementation of the BSC and CIPP evaluation models in organizations.

2. Success Case Method. The Success Case Method was employed as an evaluation tool to analyze each of the case studies presented in this particular study by finding out how well the BSC and CIPP models have worked in those organizations that have implemented it successfully.

3. Sample. Case studies came from different sources and organizations from a range of industries, and covered different issues concerning BSC and CIPP implementation process and results (i.e., BSC and CIPP design issues, implementation process and principles, integration of different models such as BSC and Six Sigma, use of elements (i.e., strategy maps, use of indices), BSC software), link between BSC and compensation. The number of case studies including BSC and CIPP evaluation models were six.

4. Data preparation. The data was prepared for analysis using a four step approach and case studies content information was organized to include: (a) general information and background about each organization; (b) problem statement and rationale for implementing either the BSC or
CIPP models; (c) information regarding the implementation process; (c) information on products or outcomes in each of the case studies reviewed.

5. Analysis. The data was analyzed by using the five step SCM method that included a case study protocol and by seeking patterns and themes in data and cross comparison.
CHAPTER IV
RESULTS

This study reviewed the different evaluation models that were used for strategic decision-making in organizations as the context for answering the following research questions:

2. What are the similarities and differences related to actual implementation of BSC and CIPP Models in terms of its methods, including: evaluation components, evaluation indicators, data collected to support the evaluation, evaluation implementation protocol, and qualitative and quantitative analyses?

3. What are the outcomes of these two (BSC and CIPP) evaluation models; what are the similarities and differences?

4. What are the critical factors that seem to be associated with successful applications of the BSC and CIPP Model?

As presented in Chapter II the first research question, what are the differences and interrelationships among the BSC, CIPP, TQM, Six Sigma, and AOP evaluation models was addressed. As described in Chapter I and Chapter II, the BSC, CIPP, TQM, Six sigma, and AOP evaluation models have been developed and implemented in many organizations, and all were used to implement strategic performance evaluation that facilitate managers strategic decision-making, planning and control.

The results of this study are presented as follows. First, a review of the similarities and differences related to the implementation of BSC and CIPP Model in terms of its methods are described and discussed in the context of six case studies.
The discussion of these results addresses the methods that each evaluation model used, including: evaluation components, evaluation indicators, data collected to support the evaluation, implementation protocol, and analysis of both qualitative and quantitative information. Summary tables were used to illustrate the comparisons of methods used in each evaluation model. Second, the outcomes of BSC and CIPP evaluation models, their similarities and differences are presented and discussed. A summary table was used to illustrate comparisons of outcomes that were obtained in each evaluation model. Finally, critical factors associated with successful applications of the BSC and CIPP Model were identified and discussed.

Research Question # 2

What are the similarities and differences related to the implementation of BSC and CIPP Model in terms of its methods, including: evaluation components, evaluation indicators, data collected to support the evaluation, implementation protocol, and analysis of both qualitative and quantitative information?

Six case studies, four BSC and two CIPP formed the context for answering the second research question. Final evaluation reports, (The Spirit of Consuelo, 2002; and Nasa-AESP, 2004) and published case studies reports (Siemens AG, 2004; Hilton Hotel Corporation, 2000; Mobil North America, 2000; United Parcel Services, 1999), provided the necessary data. Results are presented across six tables: evaluation components, evaluation indicators, data collection methods, implementation protocol, and qualitative and quantitative analyses.
Evaluation Components

Throughout this dissertation, the term “evaluation components” refers to the performance and evaluation elements involved in each model (BSC, CIPP). These elements provide evaluative information that is used as a guide to assist managers and practitioners in the implementation process of these evaluation models that is used to inform and to improve both strategic and operational management decisions.

Evaluation components involved in the BSC model include the following: customer, financial, internal, and learning and growth. Evaluation components involved in the CIPP model include the following: context, input, process, and product. Product evaluation may be divided into impact, effectiveness, transportability, and sustainability evaluation components in order to assess long-term outcomes.

Table 5.

Evaluation components used in BSC and CIPP evaluations

<table>
<thead>
<tr>
<th>BSC CASE # 1 SIEMENS AG</th>
<th>Evaluation Components:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼Evaluation components included in the BSC Model included measures on the four different BSC perspectives: Customer, Financial, Internal, and Innovation and Learning. The senior management team at Siemens identified the BSC as an effective tool for simplifying strategy implementation and reconnecting strategic direction with operational activities. The BSC model was linked with six sigma. Additionally, the BSC was selected as an opportunity to drive up market share. Siemens adopted the BSC model in 1999.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 2 HILTON</th>
<th>Evaluation Components:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼The BSC process of implementation at Hilton Hotels included measures on the different perspectives: Operational Effectiveness, Revenue maximization, Loyalty, Brand Standards, and Learning and Growth. The BSC was implemented to drive up operational performance and customer satisfaction at each hotel. In addition, the BSC model linked many customer TQM initiatives that included important performance indicators into a single, focused, strategic direction throughout the organization. Hilton adopted the BSC model in 1997.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 3 MOBIL</th>
<th>Evaluation Components:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼Mobil included a Customer, Financial, Internal, and Learning and Growth perspectives/elements in their BSC implementation process. The management team at Mobil used the BSC model to mapped-out a two-way customer-focused strategy for generating higher volume on premium-priced products and services, while reducing costs and improving productivity.</td>
<td></td>
</tr>
</tbody>
</table>
Table 5 illustrates the interrelatedness of some of the evaluation components that are used in both the BSC and CIPP models. When analyzing the BSC and CIPP evaluation models it is apparent that both models have compatible evaluation components. For instance, evaluation components included in the BSC Model (Kaplan and Norton, 1992) included measures representing the four different BSC perspectives: Customer, Financial, Internal, and Innovation and Learning, with some variations in the selection of these perspectives depending on the uniqueness of each organization’s strategy. On the other hand, evaluation components under the CIPP case studies included measures representing the four different types of evaluation components.
(context, input, process, and product). Additionally, Nasa case study included the four types of evaluation as well as some elements of the Scriven (1967) “formative-summative” evaluation approach.

Both the BSC and CIPP models use similar evaluation components to provide information to inform decision-making and accountability concerning the different components. For instance: the BSC helped managers to implement strategy by translating the vision and strategy into a set of operational objectives across each perspective and down through all levels of the organization. The organization's activities, resources, and initiatives were aligned to the organization’s strategy. The CIPP model provided evaluators with a comprehensive description and assessment of how context (goals), inputs (resources), processes (actions/activities), and products (intended and unintended outcomes) across these organizations were managed and deployed for understanding, improvement, and accountability purposes.

**Evaluation Indicators**

Throughout this dissertation, the term “evaluation indicators” refers to the performance and evaluation metrics or measures involved in each evaluation component in the BSC model (customer, financial, internal, and learning and growth), and the CIPP model (context, input, process, product). These evaluation indicators were used to improve an organization’s measurement system by helping managers and evaluators to develop relevant indicators that were focused on the critical factors included in the strategic plan of the organization (in case of the BSC model) or in the evaluation project (in case of the CIPP model).
### Table 6.

**Evaluation indicators used in BSC and CIPP evaluations**

<table>
<thead>
<tr>
<th>BSC CASE # 1</th>
<th>SIEMENS AG</th>
<th>Evaluation Indicators:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼After an assessment including the business’s values, environment, strengths and weaknesses; metrics were developed focused on three critical success factors, measuring:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.- Speed- defined as clear and fast processes, logistics excellence, and time-to-market.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.- Innovation- defined as smart ideas and courageous visions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.- Volume- global presence, brand awareness/image and technological excellence.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ Process improvement objectives are supported by strategically important Key Performance Indicators (KPIs).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ Some KPI, such as cost of non-conformance was common across business unit scorecards. Non-conformance costs were a central element of the scorecard.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 2</th>
<th>HILTON</th>
<th>Evaluation Indicators:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼ Hilton executives selected the following value drivers (or corporate strategic direction) that drive value for the organization: operational effectiveness, revenue maximization, loyalty, brand standards, and learning and growth.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ From the value drivers, Hilton executives selected the key performance indicators (KPIs), that represented the property specific goals that each hotel was to achieve.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ Every hotel within the organization were focused in the same “value drivers”; however, their KPIs were unique to their property, and were viewed as an integrated performance team.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ A measures under the Operational Effectiveness value driver, was: Cash Flow/GOP/ Flow-thru.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ Measures under the Revenue Maximization value driver, were: Room RevPar, and Market Share.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ A measure under the Loyalty value driver, was: Customer/Team Member.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ A measure under the Brand Management value driver, was: Brand Standards.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ Measures under the Learning and Growth value driver, were: Training/Orientation/ Diversity; and Skills Certification.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 3</th>
<th>MOBIL</th>
<th>Evaluation Indicators:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼ Relevant evaluation indicators were selected in the four BSC perspectives (customer, financial, internal, and learning and growth) to measure the progress and impact on Mobil NAM&amp;R’s customers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ One strategic theme under the Financial Perspective was identified, and focused on financial growth of Mobil NAM&amp;R and included the following strategic objectives: return on capital employed; existing asset utilization; profitability; industry cost leader; and profitable growth.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ Two strategic themes under the Customer Perspective were identified and they focused on delighting the customer and win-win dealer relationships. Some of the strategic objectives, included: continually delight the targeted consumer; build win-win in relations with dealer.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▼ Five strategic themes under the Internal Perspective were identified, and they focused on build the franchise, safe and reliable, competitive supplier, quality, and good neighbor. The strategic objectives, included: innovative products and services; best in-class franchise teams; refinery performance; inventory management; industry cost leader; on specification-on time orders; improve environmental safety.</td>
</tr>
</tbody>
</table>
Table 6-Continued

<table>
<thead>
<tr>
<th>BSC CASE # 3 MOBIL</th>
<th>teams; refinery performance; inventory management; industry cost leader; on specification- on time orders; improve environmental safety.</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼One strategic theme under the Learning and Growth Perspective was identified, and it focused on motivated and prepared Workforce, it also included the following strategic objectives: climate for action; core competencies and skills; access to strategic information.</td>
<td></td>
</tr>
<tr>
<td>▼ Examples of the evaluation indicators selected under the financial perspective were: return on capital employed (ROCE); cash flow; net margin rank (vs. competition); full cost per gallon delivered (vs. competition); volume growth rate and industry; premium ratio; and non-gasoline revenue and margin.</td>
<td></td>
</tr>
<tr>
<td>▼ Examples of the evaluation indicators selected under the customer perspective were: share of segment in selected key markets; shopping rating; dealer gross profit growth; and dealer survey.</td>
<td></td>
</tr>
<tr>
<td>▼ Examples of the evaluation indicators selected under the internal perspective were: new product return on investment (ROI); new product acceptance rate; dealer quality score; yield gap; unplanned downtime; inventory levels; run-out rate; activity cost vs. competition; perfect orders, number of environmental incidents; and days away from work rate.</td>
<td></td>
</tr>
<tr>
<td>▼ Examples of the evaluation indicators selected under the Learning &amp; Growth perspective were: employee survey; personal BSC (%); strategic competency availability; and strategic information availability.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 4 UPS</th>
<th>Evaluation Indicators:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼ UPS executives selected four “Point of Arrival” (POAs) evaluation indicators that represented the essence of their strategic levers for success. These POA indicators were: 1.- Customer Satisfaction Index 2.- Employee Relations Index 3.- Competitive Position; and 4.- Time in Transit.</td>
<td></td>
</tr>
<tr>
<td>▼ The BSC measurement system focused on “results tracking” rather than “activity tracking”, where each strategic measure must connect analytically (cause-and-effect) with one or more of the POA goals.</td>
<td></td>
</tr>
<tr>
<td>▼ Examples of indicators that were used under the Customer Satisfaction perspective include: claims index, concerns index, data integrity, and percent (%) of package-level detail.</td>
<td></td>
</tr>
<tr>
<td>▼ Examples of indicators that were used under the Financial perspective include: Volume/revenue/cost index; Profit index.</td>
<td></td>
</tr>
<tr>
<td>▼ Examples of indicators that were used under the Internal perspective include: Quality report card; Operations report card.</td>
<td></td>
</tr>
<tr>
<td>▼ Examples of indicators that were used under the People (Learning and Growth) perspective include: Safety; Employee retention; Employee Relations index.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CIPP CASE # 1 SPIRIT OF CONSUELO</th>
<th>Evaluation Indicators:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼ Relevant indicators were selected in the four types of evaluation (context, input, process, and product) to measure progress and impact of this project on the quality of life for children and families.</td>
<td></td>
</tr>
</tbody>
</table>
### CIPP CASE # 1 SPIRIT OF CONSUELO

Examples of Family Quality of Life and Children’s Quality of Life Indicators covered areas/issues such as employment, health, safety, housing, family structure, safety, education, community. Indicators used under these different areas were: annual family income, available quality health care services for all family members, evidence of drug or alcohol abuse, appropriate size and design of the house to accommodate the needs of those who inhabit it, total number of persons living in the home, educational level of family members.

Examples of process evaluation indicators to assist and assess project implementation, included: assessment of project moving toward anticipated outcomes (review of annual work plan), resources appropriately directed to fulfill project goals (review of project documentation/interviews), documentation of activities, evaluation feedback, internal evaluation processes.

Examples of impact evaluation indicators included: intended beneficiaries that have been reached and identified needs been met through the conduct of needs assessment compared with earlier assessments, interviews, records of the project, and assessment of project’s impact on the community through interviews with community leaders, inspection of the project site and the larger community.

### CIPP CASE # 2 NASA

#### Evaluation Indicators:

A Delphi survey was conducted to select those evaluation indicators or outcomes that resulted from the work of AESP members. Indicators were categorized under four different client groups, including: student groups, classroom teacher groups, administrators groups, and professional education groups), that have received AESP services, and were used to measure the impact of AESP activities on these groups.

Examples of indicators that were used in this project under the student groups are: feedback letters from teachers with student assessment and descriptions of student behavior, presence of newspaper articles/media clips about positive reaction to NASA-AES visit, results of local evaluation efforts, application and participation in other NASA programs, students use of real time NASA information.

Examples of indicators that were used in this study under the classroom teacher groups are: feedback indicating use and value of activity, demonstrate use of modeled teaching strategies/techniques, perceive value in and seek NASA certification to use and teach NASA related content/materials, sample lessons showing NASA products in use.

Examples of indicators that were used in this study under the administrator groups are: feedback outlining implementation of workshop materials, assessment of needs to implement curriculum aligned with state standards/expectations, partnerships to enhance professional development opportunities for teachers.

Examples of indicators that were used in this study under the professional education groups are: requests for additional information on resources, statements indicating use of materials, requests for presentations or other assistance.

Examples of indicators that were used in this study under the State Department of Education are: development of organizational support groups that link private and public sector stakeholders, aligned NASA products with state standards of learning, awareness of NASA materials and active sharing of information with others with an interest or need to know.

Other indicators that were used in this study are: greater availability, access to, and use of aerospace materials and activities among organizations; facilitates collaborative actions with interested stakeholders from various occupations/professions to advance understanding of aerospace science.
Table 6 illustrates the different evaluation indicators that were used in both models (BSC and CIPP). The value of these evaluation indicators that were used in those organizations that implemented the BSC and CIPP models, is that they provided managers and evaluators with information in order to identify those actions that should be taken in order to: accurately reflect the organization's performance current situation, guide employees to take the right decisions in situations where action is required, and determine the effectiveness of those actions.

The different organizations included in this study, employed BSC and CIPP’s indicators as a means to develop close loop feedback systems that embodied situational analysis of information, corrective actions, and result evaluation.

In the case of the BSC, the use of indicators under the four different evaluation components (customer, financial, internal, learning and growth) was seen by managers as an effective way to measure the progress and impact on the organization’s strategic plan. The BSC was used as a comprehensive evaluation model for defining and communicating strategy, for translating this strategy to operational terms, and for measuring the effectiveness of the organization’s strategy implementation.

The CIPP evaluation model provided managers with indicators under the four types of evaluation (context, input, process, and product), in order to measure progress and impact of the project under study. In both CIPP case studies (The Spirit of Consuelo and Nasa), indicators were selected and categorized under the different clients or stakeholders’ groups.
It should be noted, that in the particular CIPP model’s case studies that were included in this study, the methods that were used focused on qualitative information; therefore the indicators that were employed were also inclined to be mostly qualitative. This could be explained as a result of the projects’ scope assessing the effects of the projects on people’s perceptions, attitudes, and behavior changes. It is important to emphasize, that in these specific types of projects that use qualitative indicators, a critical challenge derives from the interpretation of those qualitative measures. There is a premium on the evaluator’s ability to clearly articulate measures and develop instruments in such a way that the project’s interpretations of measures will vary as little as possible. The aim when creating and using indicators is to provide an objective view of an organization’s performance, and thus subjectivity should be avoided. This specific problem was clearly addressed in the Spirit of Consuelo and NASA’s case studies by using a triangulation method were different data sources and procedures were used to get a fix on a particular measurement issue.

Data Collected to Support BSC and CIPP Evaluations

Throughout this dissertation, the term “data collected to support BSC and CIPP evaluations” refers to the information collected from each case study included in this study regarding the sampling or targeted population (i.e., business units or clients groups) that were included. Additionally, information related to the context and local or organizational environment under which either the BSC or CIPP was implemented is provided.
Data collected to support BSC and CIPP evaluations

| BSC CASE # 1 | SIEMENS AG | Data Collected:  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼ The implementation of the BSC started at a business unit (SBU), including the senior management group (29,000 person business unit) from a total population of 114,000 employees. The BSC was not implemented in a conventional top-down cascade, but alignment of measures was present across business unit scorecards.</td>
</tr>
</tbody>
</table>

| BSC CASE # 2 | HILTON | Data Collected:  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼ The implementation of the BSC began at the Hotel Operations level of Hilton Hotels Corporation. More specific information regarding the sampling and population, was not disclosed in this case study.</td>
</tr>
</tbody>
</table>

| BSC CASE # 3 | MOBIL | Data Collected:  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼ The implementation of the BSC began at Mobil’s North America Marketing and Refining Division (NAM&amp;R). Mobil NAM&amp;R, implemented its BSC into its 18 business units and 14 strategic partners (freestanding service companies called “servcos”).</td>
</tr>
</tbody>
</table>

| BSC CASE # 4 | UPS | Data Collected:  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼ The implementation of the BSC initiated at the corporate level, including UPS executives, and then cascading the BSC first to each region and district, and then to each business unit and individual level. The BSC was implemented in a conventional top-down cascade process. The BSC was implemented in 11 domestic and 5 international regions, comprising 60 districts and 1600 business units in the United States. UPS has a total population of 326,800 employees in more than 2000 countries.</td>
</tr>
</tbody>
</table>

| CIPP CASE # 1 | SPIRIT OF CONSUELO | Data Collected:  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼ The Ke Aka Ho’ona evaluation project used the CIPP model to support 79 low income families to construct their own houses and to develop a healthy, values-based community in the Oahu’s Waianae coast, one of Hawaii’s most depressed and crime-ridden areas.</td>
</tr>
</tbody>
</table>

| CIPP CASE # 2 | NASA | Data Collected:  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>▼ The evaluation of the Aerospace Education Services Program (AESP) program was an agreement between the Oklahoma State University (OSU) and the National Aeronautics and Space Administration (NASA). The target population for this project were divided in two groups: providers and clients. Providers included: NASA HQ education administrators, OSU-AESP management team, NASA field center precollege officers, AESP specialists and administrative assistants. Clients included: elementary and secondary students and teachers plus local contact persons, State-level science curriculum coordinators and leaders, other educators and local officials representing professional education organizations, higher education teacher preparation programs that participated in or received services from AESP.</td>
</tr>
</tbody>
</table>

This table illustrates the general context under which each BSC or CIPP evaluation model were implemented at these different organizations. The review of published BSC case studies showed the initial scorecard process usually initiated in a strategic business unit (SBU). The SBU selected by each organization was one that
conducted activities across an entire value chain, including: innovation, operations, marketing, selling, and service. Concerning the CIPP’s case studies, results showed that depending on the evaluation project’s goals and objectives the CIPP evaluation process started at the particular target audiences or groups that were identified by the client.

**Evaluation Implementation Protocol used in BSC and CIPP Evaluations**

Throughout this dissertation, the term “evaluation implementation protocol” refers to the similarities and differences related to the implementation process of either the BSC or CIPP evaluation models. Implementation protocol provides information on how each evaluation model have been used to measure and manage an organization’s performance in the context of their utility for managerial decision making.

In addition, information regarding each model implementation’s management principles and approaches (top-down or decentralized approach), and time requirements is also provided.

Table 8.

**Evaluation implementation protocol used in BSC and CIPP evaluations**

<table>
<thead>
<tr>
<th>BSC CASE # 1 SIEMENS AG</th>
<th>Evaluation implementation protocol:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼ The BSC was first implemented at the business unit level (Siemens IC Mobile).</td>
<td></td>
</tr>
<tr>
<td>▼ The BSC was integrated with Siemens’ Six Sigma program in a one-year, four phase program from strategy formulation to operation.</td>
<td></td>
</tr>
<tr>
<td>▼ The management group selected the BSC model as a means to simplify strategy implementation and to reconnect Siemens’s strategic direction with operational activities.</td>
<td></td>
</tr>
<tr>
<td>▼ The BSC implementation at Siemens, included a four-stage phase process:</td>
<td></td>
</tr>
<tr>
<td>1. Business Assessment</td>
<td></td>
</tr>
<tr>
<td>2. Development of business strategy</td>
<td></td>
</tr>
<tr>
<td>3. Operationalize business strategy</td>
<td></td>
</tr>
<tr>
<td>4. Operations</td>
<td></td>
</tr>
<tr>
<td>▼ Business assessment, under this phase Siemens’ management group used a tool “value chain” to assess their business, including: value disciplines, environment, strengths,</td>
<td></td>
</tr>
</tbody>
</table>
| BSC CASE # 1 | SIEMENS AG | opportunities, and weaknesses areas.  
|             |           | ▼ Development of business strategy phase, once the business was assessed, the management group developed the business strategy, and end up with some |
| BSC CASE # 2 | HILTON | Evaluation implementation protocol:  
|             |           | ▼ The BSC was deployed at Hilton in a three-year, five phase program, including all five principles of a Strategy-Focused Organization (Kaplan & Norton, 1999).  
|             |           | ▼ In addition, the BSC Model was used as a means to implement corporate strategy and goals into every Hilton hotel. The BSC was used not only as a model to integrate all performance measures and the various TQM and change initiatives; but it was applied also to improve operating and customer results.  
|             |           | ▼ The BSC’s implementation process started at the Hilton Hotel Operations business unit level. The executive and operations teams defined the vision, value drivers (those areas of strategic importance that drives value for Hilton throughout the organization), the key performance indicators (the specific metrics at the strategic business unit level that were quantified, goals set, and results measured against those KPIs, and the five constituencies or Hilton’s stakeholders, including: Customers-External (guests), Customers-Internal (team members), Company Shareholders, Corporate Strategic Partners, and Community.  
|             |           | ▼ After the BSC Model was implemented at each hotel, it was then cascaded to area and regional vice-presidents, to department managers, and individual team members. By doing this, Hilton’s employees at all levels were not only aligned to the corporate strategic direction, but also they were compensated based on their own performance KPIs within their control. |
| BSC CASE # 3 | MOBIL | Evaluation implementation protocol:  
|             |           | ▼ The BSC was implemented at Mobil in a five-year (1994-1999), five phase program and applied all five principles of a Strategy-Focused Organization2 (Kaplan & Norton, 1999), including:  
|             |           | 1.- Mobilize Change through Executive Leadership  
|             |           | 2.- Translate the strategy to operational terms  
|             |           | 3.- Align the organization to the strategy  
|             |           | 4.- Make strategy everyone’s job  
|             |           | 5.- Make strategy a continual process.  
|             |           | ▼ Mobilize change through executive leadership, was used during the BSC implementation process to gain committed ownership from the executive team that was composed by finance, operations, information technology, and human resources team members. The team members became accountable for developing a vision and for various pieces of the strategy.  
|             |           | ▼ Translate strategy into operational terms, was used by applying the strategy map tool to translate its strategy into strategic objectives, and then into measures under the four BSC perspectives. In addition, Mobil’s strategy BSC (or corporate scorecard) was designed as an “Guide” for the development of the 18 business units balanced scorecards that were created after the corporate scorecard.  
|             |           | ▼ Align the organization to the strategy, was used to achieve strategic alignment by focusing Mobil NAM&R division in those strategic themes and priorities defined in their strategy map in the corporate and business unit levels. Each business unit chose those objectives that support the corporate scorecard. Six major strategic objectives guided the scorecard development, consisting of: |

---

2 A Strategy Focused Organization, places strategy at the center of its management processes — strategy is central to its agenda (Kaplan & Norton, 1999).
Table 8-Continued

<table>
<thead>
<tr>
<th>BSC CASE # 3</th>
<th>MOBIL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. Achieve financial returns (as measured by ROCE—return on capital employed)</td>
</tr>
<tr>
<td></td>
<td>2. Delight targeted customers with a great food and fuel-buying experience</td>
</tr>
<tr>
<td></td>
<td>3. Develop win-win relationships with dealers and retailers</td>
</tr>
<tr>
<td></td>
<td>4. Improve critical internal processes—low cost, zero defects, on-time deliveries.</td>
</tr>
<tr>
<td></td>
<td>5. Reduce environmental, safety, and other health-threatening incidents</td>
</tr>
<tr>
<td></td>
<td>6. Improve employee morale</td>
</tr>
<tr>
<td>▼</td>
<td>Making strategy everyone’s job, was used to connect the Division’s strategic goals to individual activities and performance. Mobil linked compensation to scorecard-based outcomes, where employees were able to set up personal work objectives that were aligned with the corporate scorecard, and then they were rewarded for both individual and team accomplishment.</td>
</tr>
<tr>
<td>▼</td>
<td>Making strategy a continual process, was used to linked strategy to the budgeting process (yielded both operational and strategic budgets), to the management meeting (yielded both operational and strategic performance reviews), and to the learning process (yielded both operational and strategic information systems).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 4</th>
<th>UPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evaluation implementation protocol:</td>
</tr>
<tr>
<td>▼</td>
<td>The BSC was deployed at UPS in a three-year, five phase program that included a results-driven measurement system that focused employees at all levels on customers and solutions.</td>
</tr>
<tr>
<td>▼</td>
<td>In addition, the BSC implementation also focused on the definition and measurement of results rather than activities, and the BSC was implemented within the context of UPS existing Total Quality Management system.</td>
</tr>
<tr>
<td>▼</td>
<td>The BSC implementation steps included:</td>
</tr>
<tr>
<td>1. Educating senior management about Total Quality principles</td>
<td></td>
</tr>
<tr>
<td>2. Establishing “Point of Arrival” (POA) goals at the Corporate level</td>
<td></td>
</tr>
<tr>
<td>3. Establishing a BSC business plan with baselines and POA targets for each region and district</td>
<td></td>
</tr>
<tr>
<td>4. Deploying scorecard-based plans through a Quality Improvement process (QIP) at the business unit level and a Quality Performance Review (QPR) at the individual level.</td>
<td></td>
</tr>
<tr>
<td>▼</td>
<td>The customer satisfaction perspective was used to capture the ability of the organization to provide quality services, effective delivery, and overall customer satisfaction.</td>
</tr>
<tr>
<td>▼</td>
<td>The financial perspective was used as a guide to select financial objectives that represented long-range targets.</td>
</tr>
<tr>
<td>▼</td>
<td>The innovation perspective was used to provide data regarding process results against evaluation indicators that lead to financial success and satisfied customers. In addition, it was used as a guide for choice of objectives and for identifying the key business processes at which UPS must excel. Key processes were monitored to ensure that outcomes were satisfactory. The innovation perspective provided the mechanisms through which performance expectations were achieved.</td>
</tr>
<tr>
<td>▼</td>
<td>People (learning and growth) perspective was used to capture the ability of UPS’s employees, information systems, and organizational alignment to manage the business and adapt to the BSC change. Employees at UPS were motivated, and supplied with accurate and timely information to make decisions about improving processes.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CIPP CASE # 1</th>
<th>SPIRIT OF CONSUELO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evaluation implementation protocol:</td>
</tr>
<tr>
<td>▼</td>
<td>The CIPP model included questions derived from the types of evaluation: context, input, process, product, impact, effectiveness, sustainability, and transportability and guided this evaluation project.</td>
</tr>
<tr>
<td>▼</td>
<td>Context evaluation was used in this evaluation project for goal-setting purposes, and</td>
</tr>
<tr>
<td>CIPP CASE #1</td>
<td>SPIRIT OF CONSUELO</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>helped the evaluation team to determine the target population and to clarify and update the project's goals to assure that they properly address assessed needs of the Waianae community. It also helped the evaluation team to assess the significance of outcomes through ongoing assessments of the Waianae's housing and community needs.</td>
<td></td>
</tr>
<tr>
<td>▼Input evaluation was used in this evaluation project for planning purposes, and helped the evaluation team to assure that the project's initial strategy was economically, socially feasible for meeting the assessed needs of the Waianae community.</td>
<td></td>
</tr>
<tr>
<td>▼Process evaluation was used in this evaluation project for managing purposes, and helped the evaluation team to strengthen project implementation and design in areas of operational deficiency, and helped them to maintain a record of the project's process and costs, and to report on the project's progress.</td>
<td></td>
</tr>
<tr>
<td>▼Impact evaluation was used in this evaluation project for controlling purposes, and helped the evaluation team to determine the extent to which this project reached the beneficiaries of the Waianae community.</td>
<td></td>
</tr>
<tr>
<td>▼Effectiveness evaluation was used in this evaluation project for assuring quality, and helped the evaluation team to determine the project's effects on the quality and life conditions of the Waianae community.</td>
<td></td>
</tr>
<tr>
<td>▼Sustainability and transportability evaluation were used in this evaluation project for institutionalizing disseminating purposes, and helped the evaluation team to estimate the extent to which successful aspects of this project could be sustain and applied in other settings. It also aided the evaluation team to make a bottom-line assessment of the success and significance of this project.</td>
<td></td>
</tr>
<tr>
<td>▼Formative evaluation was used for project improvement by providing annual reports containing feedback from beneficiaries and the Foundation staff, as well as the evaluator perspectives on the project's environmental factors, documented project operations, identified strengths and weaknesses. Summative evaluation was used to assess the project's success in terms of meeting the Waianae community's needs.</td>
<td></td>
</tr>
<tr>
<td>▼The use of a project development cycle, that included: 1. Project identification goal setting, in which needs from the Waianae community were identified. Under this stage the evaluation project team collected background information to choose those projects that were economically, legally, and politically feasible. 2. Project planning, involved the development project objectives that focused directly on the target group's assessed needs, and also detailed plans for those projects that were chosen to address all aspects of the project (including, technical, economic, financial, social, and institutional). Under this stage, the intent was to design and operationalize the best method(s) for meeting the Waianae community's needs. 3. Project implementation, involved constant monitoring, administration, and improvement of processes. 4. Project control, was used to assure that project services were reaching the target population. In this evaluation, the target audience needed to be redefined, because originally the Foundation hoped to use self-help housing to serve the housing needs of Hawaii's poorest of the poor. However, such persons could not qualify for the required home mortgages. It redefined its target audience to include low-income persons who could qualify for a mortgage but who would be unlikely to own a home without special assistance. Assessed needs of intended beneficiaries should have been identified and examined, before selecting an appropriate project strategy. 5. Quality assurance, was achieved by constantly monitoring both process and product to assure that valuable outcomes were achieved. This project's signs of quality and constant improvement included effective safety training; improvements in materials and their delivery; improvements in house designs; rigorous inspections of all aspects of construction; continuous on-site monitoring of the construction process by a Foundation staff member; regular external program evaluations.</td>
<td></td>
</tr>
</tbody>
</table>
6. Institutionalization/Dissemination, sought to disseminate the lessons learned in this project, including reporting on its successes and failures. It described the project's approach and setting, its beneficiaries, its strengths and weaknesses.

- The evaluation purposes were: improvement, it provided information to help project staff assess and improve the ongoing process; accountability, it helped the Foundation maintain an accountability record and keep them appraised of the project's performance in carrying out planned procedures; understanding, it helped analyze the project's background, process, and outcomes; and dissemination, it helped to inform developers and other groups about this project's mission, objectives, structure, process, and outcomes.

<table>
<thead>
<tr>
<th>CIPP</th>
<th>Evaluation implementation protocol:</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASE # 2 NASA</td>
<td>▼ An adaptation of the CIPP (Context, Input, Process, and Product) model was used, as well as, some elements of the Scriven (1967) &quot;formative-summative&quot; evaluation approach, and some basic recommendations for government-sponsored evaluation of educational programs.</td>
</tr>
<tr>
<td></td>
<td>▼ To gain additional insights into project outcomes, this evaluation focused primarily on the product/impact evaluation component in order to assess the effectiveness and impact of the Aerospace Education Services Program (AESP).</td>
</tr>
<tr>
<td></td>
<td>▼ The CIPP Model was chosen as a means to obtain timely and credible information for purposes of decision making (formative evaluation) and accountability (summative evaluation).</td>
</tr>
<tr>
<td></td>
<td>▼ The evaluation program was conducted in four major phases of effort over a three-year period. The four major program areas included: support of systemic improvement, teacher preparation and enhancement that support systemic reform, student support, and curriculum support and dissemination.</td>
</tr>
<tr>
<td></td>
<td>▼ Context evaluation was used as a guide for choice of objectives and assignment of priorities, and to maintain a record of objectives, record of needs, opportunities, and problems.</td>
</tr>
<tr>
<td></td>
<td>▼ Input evaluation, was used as a guide for choice of program strategy and input for specification of procedural design.</td>
</tr>
<tr>
<td></td>
<td>▼ Process evaluation, was used as a guide for project implementation to monitor, maintain records, and to provide periodic progress reports.</td>
</tr>
<tr>
<td></td>
<td>▼ Product Evaluation (Impact, Effectiveness, and Viability), was used to guide decisions regarding termination, continuation, or modification of the program, as well as, to maintain a record of attainments, to compare assessments of needs and costs, and for recycling decisions.</td>
</tr>
<tr>
<td></td>
<td>▼ The evaluation work centered around 19 critical questions across 5 major categories, including: program design and management, support of systemic improvement, teacher preparation programs and enhancement programs that support systemic reform, student support, and curriculum and dissemination.</td>
</tr>
<tr>
<td></td>
<td>▼ The evaluation purposes were: First, improvement by providing support to improve state-based educational reforms for NASA-AESP programs. Second, accountability by developing and maintaining evaluation indicators utilizing NASA's on-line Education Computer Aided Tracking System (EDCATS) to evaluate and monitor program results. Finally, understanding by helping to assess the effectiveness and impact of the national space education program.</td>
</tr>
</tbody>
</table>
This table presents some of the evaluation implementation protocols used in BSC and CIPP evaluation models by managers and evaluators. The review of published BSC case studies showed how adopters of the BSC model followed an implementation protocol based on five management principles of a “Strategy-Focused Organization” (Kaplan & Norton, 1999). The first principle included translating the organization’s strategy to operational terms (where evaluation indicators were defined for strategic objectives that comprised the organization’s BSC). The second principle entailed aligning the organization to the strategy (where all levels of the organization focused on strategic themes and priorities as defined by their strategy map and corresponding balanced scorecards for the corporate, business unit and support units). The third principle involved motivating employees to make strategy everyone’s job (where individuals set personal work objectives which aligned with the organization’s BSC, and were rewarded with compensation and recognition for both individual and team accomplishment). The fourth principle consisted of adapting employees to the new strategy as a continual process (where managers linked the organization’s strategy to the budgeting process, management, and learning processes). The final principle comprised mobilizing change in the organization through executive leadership (where managers mobilized change by demonstrating committed support to implement the BSC).

Moreover, it is important to note that depending on the organization’s culture the BSC implementation process followed two approaches: First a top-down implementation approach, where a strategy map was cascaded down and deployed at all levels of the organization to ensure strategy-aligned behavior and to make every
employee aware of the organization’s strategy. Second, a decentralized implementation approach in which an organization’s business unit or division level, were an operations teams were responsible for defining the vision, value drivers, and performance indicators. Under this approach, the strategy implementation process was not only managed by senior executives, but also by the middle-level managers and employee teams that actually perform the work.

The culture of the organization played an important role when determining which BSC implementation approach was to be followed; however, in all cases a key issue for successful BSC implementation was employee involvement.

The review of published CIPP case studies showed that adopters of the CIPP model followed an implementation protocol based on four types of evaluation (context, input, process, product) (Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985) to assess and improve their evaluation projects including: context evaluation (that was used for goal-setting purposes and helped evaluators to determine the target population and to clarify and update the project’s goals); input evaluation (that was used for planning purposes and aided evaluators to choose a program strategy); process evaluation (that was to help evaluators to strengthen project implementation and design in areas of operational deficiency) and product evaluation including in some cases an impact, effectiveness, sustainability, and transportability evaluation component (that was used for controlling, quality, and disseminating purposes. These types of evaluation helped evaluators to make decisions regarding termination, continuation, or modification of the program, as well as, to maintain a record of attainments).
Additionally, a formative and summative type of evaluation were used by managers. Formative evaluation was used for project improvement by providing annual reports containing feedback from beneficiaries of the project, as well as the evaluator perspectives on the project’s environmental factors, documented project operations, identified strengths and weaknesses. Summative evaluation was used to assess the project’s success in terms of meeting the stakeholders’ needs.

The time required to implement the BSC evaluation model, varied between 3 to 5 years, and underwent three phases: First, a mobilization phase consisting of a three-to six-month period devoted to executive-level momentum building by communicating the need for change, building the leadership team, and clarifying the vision/strategy. Second, a design and rollout phase, including a six-month period in which the new strategy was rolled out at the top levels of the organization. Third, a sustainable execution phase involving a 12-24 month period where the strategy was integrated into the day-to-day work and culture of the organization. Balanced Scorecards were used to educate and align individuals.

The CIPP model implementation’s time requirements varied according to the scope of the evaluation project. The Spirit of Consuelo case study was an eight-year evaluation study, and the Nasa case study was a three-year study. The CIPP model provided evaluators with greater flexibility in terms of the types of evaluation that were needed to be conducted depending on the previously assessed stakeholders’ needs. In the particular CIPP’s case studies that were included in this study four types of evaluation (context, input, process, and product) were included. However, not all four types of evaluation are always applied. As mentioned earlier, in order to
determine which evaluation components of the CIPP model to employ and what information to collect, the evaluator need to identify and address the client's purpose for the evaluation.

**Qualitative and Quantitative Analyses in BSC and CIPP Evaluations**

Throughout this dissertation, the term "qualitative and quantitative analyses" refers to the different tools including qualitative and quantitative methods that are used in the BSC and CIPP evaluation models. Depending on the evaluation context, managers or evaluators may employ a range of various performance and evaluation's tools and methods to plan and carry out the implementation of the BSC or CIPP evaluation models in their organizations.

Table 9.

**Qualitative and quantitative analyses in BSC and CIPP evaluations**

<table>
<thead>
<tr>
<th>BSC CASE # 1 SIEMENS AG</th>
<th>Qualitative and Quantitative Analyses:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼ Process mapping tools were employed as “strategy maps” tools to describe the logic of the linkage model, and then to translate it into the BSC’s key evaluation indicators, targets, and initiatives.</td>
<td></td>
</tr>
<tr>
<td>▼ An intranet site, was used to allow employees to review their performance data including evaluation indicators on each perspective.</td>
<td></td>
</tr>
<tr>
<td>▼ A policy deployment process was employed to ensure that employees understood the strategy, and to cascade the BSC process. Through this policy deployment process, employees from one level made commitments to specific targets with the level above. Such target agreements were a key component in order to put into operation the business strategy.</td>
<td></td>
</tr>
<tr>
<td>▼ Project management tools and cross-functional teams were also used to make the strategy operational. For instance: “Scorecard User Clubs” were used to facilitate the dissemination of scorecard learning and experiences. These user clubs included topics such as: the implication of the BSC on the budgeting process, BSC communication mechanisms.</td>
<td></td>
</tr>
<tr>
<td>▼ Management workshops were conducted involving each unit’s senior management in order to construct the BSC. These meetings were held over a 3-4 month timeframe. The objectives of conducting these workshops were to develop the process mapping (or strategy maps) tools, and then to map the interrelated strategic objectives by having executives to identify the strengths, weaknesses, opportunities, and threats (SWOT analysis), and then cluster them into themes and slot each one into the BSC perspective it best fitted.</td>
<td></td>
</tr>
<tr>
<td>▼ Strategy Matrix was used as a useful visualization and summarization tool. The strategy matrix, displayed also information about the objectives, targets, and initiatives in one table. It also pointed to areas where scorecard elements might be out of balance. Employees used this strategy matrix tool to prioritize spending for projects, where they had a cluster of initiatives around one objective, while other objectives did not have supporting...</td>
<td></td>
</tr>
</tbody>
</table>
Table 9-Continued

<table>
<thead>
<tr>
<th>BSC CASE # 1</th>
<th>SIEMENS AG</th>
</tr>
</thead>
<tbody>
<tr>
<td>initiatives. The strategy matrix reflected a strategic theme, so one matrix was prepared for each theme</td>
<td></td>
</tr>
<tr>
<td>▼ What-if analysis were used to forecast future performance. By plugging in different values from various measures, managers at Siemens were able to examine the effect on related indicators. The what-if analysis provided another opportunity to examine the assumptions made when constructing the strategy map.</td>
<td></td>
</tr>
<tr>
<td>▼ Core solutions such as merging BSC with Six Sigma, were used not only as a means to facilitate the transitions between analytic and operations tasks, but also to improve the impact of the BSC at Siemens.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 2</th>
<th>HILTON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative and Quantitative Analyses:</td>
<td></td>
</tr>
<tr>
<td>▼ Value driver tree tool (same as strategy map) was used to communicate a unified view of the overarching strategy to employees at Hilton. The value tree helped executives at Hilton to define the corporate direction and to align internal processes, strategic objectives, initiatives, key evaluation indicators, and target scores.</td>
<td></td>
</tr>
<tr>
<td>▼ BSC - web portal was used to provide hotel team members with feedback on their own individual evaluation indicators. The BSC portal also allowed team members to monitor their potential bonus as linked to their performance. Moreover, by using the BSC portal tool for measurement and information, managers at each hotel were provided with specific data that allowed them to get to the root cause of problems.</td>
<td></td>
</tr>
<tr>
<td>▼ Drill-down BSC software analytic capability, enabled employees at Hilton to determine a detailed information about the source(s) of the results, and to take steps for continuous improvement. Because the each BSC included evaluation indicators that are highly aggregated, and derived from multiple data points, this drill-down tools were needed to make improvements, and allowed employees to drill down indicators to increasingly lower levels of detail.</td>
<td></td>
</tr>
<tr>
<td>▼ Statistical analysis, including data mining and analysis computer tools were used to explore large quantities of data for comprehensive analysis to discover hidden relationships, and to uncover patterns, trends, associations, or anomalies for proactive decision making. In addition, an online analytical processing tool (OLAP), was used for multidimensional views of combination of data that were rotated for reporting and inquiries.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 3</th>
<th>MOBIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative and Quantitative Analyses:</td>
<td></td>
</tr>
<tr>
<td>▼ A strategy map was used as a tool to describe the logic of the strategy, and then managers translated it into measures and targets in order to demonstrate clearly visible links and flow in a series of strategic objectives up to ultimate objectives of financial performance. Mobil's strategy map illustrated all four BSC perspectives with cause-effect linkages.</td>
<td></td>
</tr>
<tr>
<td>▼ Methods used during the process of BSC implementation at Mobil, included: monthly BSC reports, periodic newsletters, executive presentations, customer, distributor, and market, and employees' surveys; safety index, environmental index, structure tree.</td>
<td></td>
</tr>
<tr>
<td>▼ Safety index was used to simplify data and aid in analysis. This safety index was tracked once a month and it provided the operations manager and others with an assessment of the safety level of the entire refinery. One-half of the index was based on actual accidents that occurred and their relative severity. The cost of each accident in lost time and medical expenses was also figured in. The remaining 50% of the index as a preventive measure that included safety audit scores, safety training, safety improvement actions.</td>
<td></td>
</tr>
<tr>
<td>▼ BSC reports were used to provide monthly feedback at every site, by displaying objectives and measures in the four BSC perspectives, and current and year-to-date performance.</td>
<td></td>
</tr>
<tr>
<td>▼ Scoreboard, was used to showed the critical information relative to each BSC perspective, objective, and evaluation indicator. The scoreboard was filtered either by theme or was set to</td>
<td></td>
</tr>
</tbody>
</table>
Table 9-Continued

<table>
<thead>
<tr>
<th>BSC CASE # 3 MOBIL</th>
<th>Qualitative and Quantitative Analyses:</th>
</tr>
</thead>
<tbody>
<tr>
<td>display all Mobil’s corporate indicators. Information presented in this scoreboard included: current and previous values, targets and trends of evaluation indicators, comments, and links to analysis so that an employee was able to answer questions readily. Color and graphics were used to convey whether an evaluation indicator was on target or needed attention.</td>
<td></td>
</tr>
<tr>
<td>▼ A first round of interviews were conducted by the BSC project leader with each manager to build consensus around strategic objectives. The SBU’s vision, mission, strategy, and information on the industry and competitive environment, trends in market size and growth, competitors, and customer preferences were reviewed during these interviews. The objectives of this interview were to introduce the concept of the BSC to each manager, and to get their initial input about Mobil’s strategy, and to figure out how it could be translated into objectives and evaluation indicators for the BSC.</td>
<td></td>
</tr>
<tr>
<td>▼ A first round of interviews were conducted by the BSC project leader with each manager to build consensus around strategic objectives. The SBU’s vision, mission, strategy, and information on the industry and competitive environment, trends in market size and growth, competitors, and customer preferences were reviewed during these interviews. The objectives of this interview were to introduce the concept of the BSC to each manager, and to get their initial input about Mobil’s strategy, and to figure out how it could be translated into objectives and evaluation indicators for the BSC.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 4 UPS</th>
<th>Qualitative and Quantitative Analyses:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼ No strategy map was constructed at UPS. However, it was assumed that there was a cause-effect relationship between the various evaluation indicators in their corporate scorecard and the “Point of Arrival” (POA) strategic goals that guided the company. These cause-effect relationships were discussed in the selection of the evaluation indicators, and also positive business results gave strong evidence of the intended relationships.</td>
<td></td>
</tr>
<tr>
<td>▼ Some of the methods that were used during the BSC implementation process, included communication documents/reports (i.e. quality at a glance).</td>
<td></td>
</tr>
<tr>
<td>▼ Quality at a glance was used to communicate the implementation process of the BSC program at the corporate and business unit levels. This document emphasized four key areas of focus: 1. Leadership 2. People 3. Process 4. Measurement/Assessment</td>
<td></td>
</tr>
<tr>
<td>▼ Quality improvement process was used at the business unit level for goal-setting purposes by front-line managers every six months, followed by progress review meetings on a monthly basis. It consisted on having first-line managers set these goals as targets on their business unit scorecard measures. The goals also became part of their own personal quality performance review (QPR) documents. For each of these unit goals, a strategic action plan was set in place for which the manager have personal responsibility.</td>
<td></td>
</tr>
<tr>
<td>▼ The quality improvement process’s goals represented targets set for the various BSC measures tracked by business unit. The supervisors reporting to these business unit managers were required to establish 4-6 week action plans to support those business unit goals. These action plans became part of the supervisors’ quality performance review.</td>
<td></td>
</tr>
<tr>
<td>▼ Through the use of the quality performance reviews, evaluation indicators and goals were connected to a business unit BSC. District objectives were aligned with a corporate scorecard, business unit objectives were aligned with a district scorecard, and supervisory objectives were aligned with the business unit scorecard.</td>
<td></td>
</tr>
<tr>
<td>▼ Customer satisfaction index and an employee relations index were used to simplify data</td>
<td></td>
</tr>
</tbody>
</table>
and aid in analysis. These indices included both leading and lagging indicators. “A leading or performance driver indicator is defined as a measure that indicates progress against a process or behavior. These measures are helpful in predicting the future outcome of an objective. A lag indicator is a measure that determine the outcome of an objective that indicate company performance” (Kaplan & Norton, 1992, p. 150).

### CIPP CASE # 1 SPIRIT OF CONSUELO

#### Qualitative and Quantitative Analyses:

- Multiple methods were used to gather data for each component of the evaluation. In addition, each part of the evaluation was addressed by at least three different methods (Triangulation).

  - Some of the methods that were used included:

    - Environmental analysis was used to gather contextual information by reviewing available documents and data on economics, population characteristics. It also involved interviewing persons in various roles in the area.
    - Program profiles were used as an information base with a concentration of qualitative information, including the project’s mission, goals, plan, staff, timetable, resources.
    - Traveling observers were designed and carried out to monitor and assess both project implementation and project outcomes along the way.
    - Case studies were conducted as repeated interviews with a panel of participants over time. Case studies focused on the collective perceptions of the selected families about the project and its impacts on them.
    - Interviews were conducted with the builders of the houses about three to six months after they have moved into their homes. Interviews provided information about the builders’ perceptions of the community and the process they experienced in building the houses. Interviews helped the evaluation team to understand the developing project, identify key issues related to project improvement.
    - Goal-free evaluations were conducted in year 3 and 5. This technique was useful for identifying and assessing unexpected project outcomes.
    - Feedback workshops were conducted throughout the evaluation to discuss the findings, identify areas of ambiguity in each report, and to update evaluation plans. Program personnel used these workshops to apply evaluation to their own assessments and decision processes. These workshops were held with project staff, evaluation team, project leaders and staff, and other stakeholders.
    - A synthesis-final report was used to summarize previous reports, to examine the Foundation’s documents, to gather information from Foundation staff, during the duration of this evaluation period.

  - Qualitative and Quantitative Information was collected, processed, and reported by annual written reports and a final summative evaluation report.

### CIPP CASE # 2 NASA

- In addition, an effectiveness and working conditions survey was used to determine the perceptions of the factors or conditions that influence the success of the AESP program in the field. This survey was administered to Aerospace Education Specialists (AESs) and NASA – Field Center Precollege Officers (CPOs).

  - A telephone interview with field clients of the AESP program, including: local school representatives, a person representing an informal education organization, representatives of regional or national professional education organizations, and higher education contact representatives was conducted in order to elicit responses their responses regarding the impact and ongoing influence of AESP on the local education program.

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 9-Continued

| CIPP CASE # 2 NASA | ▼ An AESP state impact survey was administered to a random sample of Council of State Science Supervisors to assess the impact of AESP services as it related to science education in the states.  
▼ Interviews with NASA-Field Center Precollege Officers (CPOs), Aerospace Education Specialists (AESs), and Aerospace Education Services program Administrative Assistant (AAs), and site visits to NASA field centers and to local schools where AESP specialists made primary presentations to students and teachers were conducted during the three years of this study.  
▼ Documents regarding the operation and performance of the AESP program were reviewed.  
▼ Database examinations were also conducted, primarily the NASA Education Evaluation and Information System.  
▼ Observations of AESP professional development activities and AESP specialist presentations were also conducted in this study.  
▼ Quantitative and Qualitative data were collected and analyzed using the SPSS for descriptive statistics (i.e., frequencies, means, standard deviations). Information from field notes of observations, interviews, focus group meetings, and from document review were transcribed and labeled to report the findings and to develop conclusions regarding the effectiveness and impact of the AESP program.  
▼ As supplements of the final evaluation report, two sub-studies were conducted the Delphi survey of indicators of AESP effectiveness; and a summary report of the 2003 survey of CPOs and AESs as a component of the evaluation of AESP. |

Table 9, illustrates the different tools and qualitative and quantitative methods that were used in the different BSC and CIPP’s case studies. It is apparent that because of the BSC evaluation model’s focus on a quantitative oriented approach the tools and methods that were used in the four BSC case studies, were related to project management and information technology tools including, strategy maps, BSC software, management workshops, data mining and computer tools to communicate BSC information.

For example, the BSC evaluation model made use of “strategy maps” (sometimes called process mapping or value driver tree tools) to describe the logic of the linkage model, and then to translate it into the BSC’s key performance indicators, measures, targets, and initiatives. The organization’s strategy is based on a set of
business processes that create value for customers, and ultimately to shareholders. The BSC evaluation model allows managers to control these processes strategically by supporting them with information, analytics, and intelligence. Strategy maps and scorecards help them describe these processes, and show how they are linked to customer and financial outcomes as well as intangible assets (i.e., employees' satisfaction).

In contrast, the CIPP evaluation model's focused on qualitative oriented tools to gather data for each evaluation component. In addition, each part of the evaluation was addressed by at least three different methods. Some of the qualitative methods and project management tools that were used in the two CIPP case studies (The Spirit of Consuelo, 2002, and Nasa, 2004) included, environmental analysis, program profiles, surveys, interviews, feedback workshops, observations, and document reviews.

For example, the CIPP evaluation model made use of “environmental analysis,” that were used to gather contextual information by reviewing available documents and data on economics, population characteristics. It also used “Effectiveness and Working conditions Surveys,” to determine the factors or conditions that influence the success of the Nasa program.

As stated earlier a key point of difference between the BSC and CIPP’s tools that were used in these specific case studies, could be explained because of the particular quantitative focus of the BSC evaluation model. The different organizations that implemented the BSC model, generated many different scorecards at the corporate, divisional, departmental, and individual level. In order to point to critical
areas that needed analysis, the balanced scorecards included measures that were highly aggregated and derived from multiple data points including details about strategies, objectives, measures, and targets. Without the right software tools, this could be a complicated and confusing process; therefore, the use of a BSC software helped managers to collect, sort, summarize, and disseminate all the data that was included in the balanced scorecards in order to communicate BSC information to employees at all levels of the organization.

The BSC and CIPP project management and quantitative and qualitative tools helped managers for the following purposes:

1. Operational Control. The emphasis on collecting data on those performance indicators that represent an organization’s strategy was to use them for assessment (learning) and predictive planning purposes. For instance a BSC-Web Portal (software), was used to provide employees with feedback on their own individual measures. A Delphi survey was used in the Nasa-Aesp’s CIPP model to assess the extent to which activities and indicators represented the effectiveness of NASA-AESP programming.

2. Assessment and Evaluation. The second purpose of collecting data was to assess what was happening in an organization and to evaluate why. The emphasis was on gaining insights and learning to better achieve the organization’s goals. Analytical capabilities were critical to improve performance by providing information that managers can use to make decisions.
3. Predictive Planning. Predictive planning, budgeting, and forecasting were used to improve an organization’s performance.

Quantitative and qualitative analytical capabilities were critical to improve performance by providing information that managers can use to make decisions.

Research Question # 3

What are the outcomes of BSC and CIPP evaluation models; what are the similarities and differences?

Throughout this dissertation, the term “outcomes” refers to some of the similarities and differences in the outcomes from both the BSC and CIPP evaluation models. Additionally, information regarding the intended and unintended outcomes is included.

Table 10, illustrate some of the similarities and differences, and intended and unintended outcomes from the BSC and CIPP evaluation models. The BSC model was found to be a strategic evaluation model for performance improvement organized across four key perspectives (customer, financial, internal, and learning and growth), which enabled an organization to articulate strategies in a set of focused, strategic objectives and measures. The BSC evaluation model was used to tell the story of an organization’s strategy in operational, measurable terms. The four BSC case studies included in this study, provided an illustration of how these different organizations turned to the BSC model in an attempt to drive organizational focus and alignment from top to bottom throughout the organization.
Table 10.

*Outcomes from BSC and CIPP evaluations*

<table>
<thead>
<tr>
<th>BSC CASE # 1</th>
<th>Outcomes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIEMENS AG</td>
<td>▼ BSC was an effective tool for simplifying strategy implementation, and reconnecting strategic direction with operational activities.</td>
</tr>
<tr>
<td></td>
<td>▼ By linking BSC and Six Sigma, Siemens’ managers were able to make the correlation between processes and non-conformance in order to reduce costs.</td>
</tr>
<tr>
<td></td>
<td>▼ Six Sigma has lead to measurable process improvement, based on figures, data, and facts.</td>
</tr>
<tr>
<td></td>
<td>▼ Sales increased 76% in one-year.</td>
</tr>
<tr>
<td></td>
<td>▼ Earnings increased 127% in one-year.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 2</th>
<th>Outcomes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>HILTON</td>
<td>▼ The BSC Model helped managers at Hilton to integrate many customer TQM/change initiatives (with different evaluation indicators) into a single, focused, strategic direction throughout the organization.</td>
</tr>
<tr>
<td></td>
<td>▼ Alignment of Hilton’s hotels (previously decentralized) with a single share vision of overall &quot;balanced&quot; performance.</td>
</tr>
<tr>
<td></td>
<td>▼ Improvements in the operations and customer satisfaction areas were achieved.</td>
</tr>
<tr>
<td></td>
<td>▼ Hilton hotel’s market revenue index increased from 104% to 106% (100% which represented a “fair share” in the local competitive market.</td>
</tr>
<tr>
<td></td>
<td>▼ Guest loyalty increased from 48% to 53%.</td>
</tr>
<tr>
<td></td>
<td>▼ From 2000 to 2002, Hilton consistently delivered a 3% higher profit margin than other full-service hotels. This 3% increase translated to a 100% increase in stock price, which were more than double that of Hilton’s competition.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 3</th>
<th>Outcomes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOBIL</td>
<td>▼ The BSC proved to be a major aligning force for the executive team, by forcing managers to narrow down their corporate strategic direction into strategic objectives.</td>
</tr>
<tr>
<td></td>
<td>▼ A 20% reduction in the cost of refining, marketing, and delivering a gallon of gasoline was achieved as a result of the BSC implementation.</td>
</tr>
<tr>
<td></td>
<td>▼ Return on capital increased from 6% to 16%.</td>
</tr>
<tr>
<td></td>
<td>▼ Volume growth exceeded industry by 2-2.5% annually.</td>
</tr>
<tr>
<td></td>
<td>▼ Improvement in cash flow was attained from $500 M/Yr to $700 M/Yr.</td>
</tr>
<tr>
<td></td>
<td>▼ Capacity utilization improved yields by $125 M per year.</td>
</tr>
<tr>
<td></td>
<td>▼ Safety incidents resulting from lost work were reduced from 150 to 30.</td>
</tr>
<tr>
<td></td>
<td>▼ Environmental incidents were reduced by 63%.</td>
</tr>
<tr>
<td></td>
<td>▼ Results from an annual employee survey showed that the impact of the BSC implementation on employees’ awareness of strategy increased from 20% to 80%.</td>
</tr>
</tbody>
</table>
### Table 10-Continued

<table>
<thead>
<tr>
<th>BSC CASE # 4 UPS</th>
<th>Outcomes:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>▼ A 9.1% increase in 1999 revenue over 1998.</td>
</tr>
<tr>
<td></td>
<td>▼ A 29% increase in operating profit compared to 1998.</td>
</tr>
<tr>
<td></td>
<td>▼ A 35.1% increase in net income in 2000.</td>
</tr>
<tr>
<td></td>
<td>▼ UPS measurement system shifted from measurement of activities to measurement of results.</td>
</tr>
<tr>
<td></td>
<td>▼ Managers were measured on results in all four BSC perspectives, not just financial.</td>
</tr>
<tr>
<td></td>
<td>▼ Linkages validated from front-line jobs to ‘Point of Arrival’ (POA) providing quantitative results. Employees were able to see the impact of their efforts on outcomes.</td>
</tr>
<tr>
<td></td>
<td>▼ Changes in measurement migrated from “rank and rating” to “criterion or threshold” systems.</td>
</tr>
<tr>
<td></td>
<td>▼ Compensation at UPS focused on a 50% of salary based on ratings from the individual’s quality performance review (QPR). This Individual QPR, were 80% based on the four BSC perspective business results connected to the manager’s business unit scorecard. The remaining 20% was derived from critical skills feedback were 50% was based on ratings by the employee’s manager and the other 50% was based on ratings by the employee and his or her peers.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CIPP CASE # 1 SPIRIT OF CONSUELO</th>
<th>Outcomes:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>▼ This project succeeded in reaching an appropriate target audience, and addressed the needs of the Waianae community, by providing people with decent, affordable housing.</td>
</tr>
<tr>
<td></td>
<td>▼ This project directly affected the lives of about 390 people, including 155 adults and 235 children.</td>
</tr>
<tr>
<td></td>
<td>▼ All features of this project were consistent with an explicit positive set of values. The values-based approach was a strong contributor to the project’s success.</td>
</tr>
<tr>
<td></td>
<td>▼ Through the participation of the beneficiaries in this project, they achieved functional homes, mortgages and land lease purchase agreements, community living guided by explicit values and rules, a safe drug-free environment, and access to a range of Foundation services.</td>
</tr>
<tr>
<td></td>
<td>▼ Children drop-out rate from schools has stopped, since families have made education a priority and the community has provided a supportive environment for children.</td>
</tr>
<tr>
<td></td>
<td>▼ Ongoing evaluation was needed at the outset of this project, throughout, and following the project. The evaluation provided a systematic, continuing look at the project and outside perspectives.</td>
</tr>
<tr>
<td></td>
<td><strong>Unintended Outcomes:</strong></td>
</tr>
<tr>
<td></td>
<td>▼ The project’s target population had to be changed because the Foundation had chosen a project strategy of self-help housing that proved unworkable with the original intended group. Assessed needs of intended beneficiaries should have been identified and examined, before selecting an appropriate project strategy.</td>
</tr>
<tr>
<td></td>
<td>▼ Some intended outcomes, such as “giving back to the larger community” were still not achieved.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CIPP CASE # 2 NASA</th>
<th>Intended Outcomes:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>▼ Overall, the impact evaluation results indicated that the NASA-AESP program was an effective logistical support unit for NASA education programming.</td>
</tr>
</tbody>
</table>
Table 10-Continued

<table>
<thead>
<tr>
<th>CIPP CASE #2 NASA</th>
<th>▼ Outcomes under the program design and management category were as follows:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. AESP’s major stakeholders reported that AESP has little impact on student science education assessment strategies.</td>
</tr>
<tr>
<td></td>
<td>2. Stakeholder respondents noted that on-site presentations made by AESP specialists to teachers and students is still the most effective strategy for delivering awareness and information about the use of NASA materials and products.</td>
</tr>
<tr>
<td></td>
<td>3. Barriers to effectively delivered the NASA message include the vast areas and formal education populations to cover by doing on-site presentations, instead of using interactive distance learning technology.</td>
</tr>
<tr>
<td></td>
<td>▼ Outcomes under the support of systemic improvement category were as follows:</td>
</tr>
<tr>
<td></td>
<td>1. The impact on and effectiveness of NASA programs and materials as core elements in state and local curriculum frameworks have been limited due to not only AESP specialist time to work intensively within any single state or local committee that is limited, but also due to the nature of AESP work.</td>
</tr>
<tr>
<td></td>
<td>▼ Outcomes under the teacher preparation and enhancement category were as follows:</td>
</tr>
<tr>
<td></td>
<td>1. AESP has not a dramatic impact on teacher preparation and enhancement programs that support systemic reform in the nation’s schools. This is a function of the vast areas and populations to be covered with 35 full-time specialists.</td>
</tr>
<tr>
<td></td>
<td>▼ Outcomes under the student support category were as follows:</td>
</tr>
<tr>
<td></td>
<td>1. The impact of AESP programming on the nation’s pre-service teacher education schools, departments, and colleges of education has been limited. The amount of time spent by AESs in these settings is quite low (1.8%).</td>
</tr>
<tr>
<td></td>
<td>▼ Outcomes under the curriculum and dissemination category were as follows:</td>
</tr>
<tr>
<td></td>
<td>1. AESP has limited exposure to and impact on informal education communities such as: museum members, planetariums, learned societies. This area of service appeared to be a low priority in a time-stretched program with many demands for service from the formal education community. In addition, rural areas of the country are not served in this sector of education.</td>
</tr>
<tr>
<td></td>
<td>Unintended Outcomes:</td>
</tr>
<tr>
<td></td>
<td>▼ Some unintended outcomes under the program design and management category that are impacting or being impacted by AESP both within and outside the agency included:</td>
</tr>
<tr>
<td></td>
<td>1. Reorganization of the NASA office of education and the elevation of education as a priority topic within the agency.</td>
</tr>
<tr>
<td></td>
<td>2. The initiation of a redesigned NASA explorer school program that requires heavy inputs of human resources and expertise from the AESP professional workforce.</td>
</tr>
</tbody>
</table>

**BSC Outcomes**

Some of the outcomes under the four different BSC perspectives that these organizations achieved after implementing the BSC model were as follows:

Under the customer perspective, managers enhanced customer satisfaction by providing lower cost products, focusing on continual quality improvement,
maximizing the organization’s operational flexibility, and enhancing customer relationships.

Under the financial perspective, managers build value in their organizations by improving its asset utilization/efficiency, decreasing its operating expenses, creating a positive cash flow, and maximizing its long-term profitability.

Under the internal perspective, managers enhance their organization’s processes by improving their organization’s manufacturing efficiency, improving the organization’s supplier management relationships, maintaining ISO certifications and enhancing community programs.

Under the learning and growth perspective, managers enhance their organization’s employees satisfaction by hiring, developing, and retaining effective performers, gathering, sharing, and using information and technology effectively, and communicating expectations, accountabilities and achievements.

The BSC was an effective tool used for simplifying strategy implementation, and reconnecting strategic direction with operational activities. The BSC proved to be a major aligning force for the executive team, by forcing managers to narrow down their corporate strategic direction into strategic objectives. The Balanced Scorecard was generally used as a performance measurement system for defining, refining and communicating strategy, for translating strategy to operational terms, and for measuring the effectiveness of strategy implementation.

CIPP Outcomes

The two CIPP case studies included in this study, also provided an illustration of how these different organizations used the CIPP model in an attempt to obtain
credible information about their programs and services for purposes of decision making and accountability. The CIPP model focused on a formative evaluation to improve these projects, as well as providing accountability records and summative evaluation reports that helped managers make informed decisions.

Some of the outcomes under the four different CIPP evaluation types that these organizations achieved after implementing the CIPP model are as follows:

Context evaluation helped evaluators to enhance their projects by: First, making decisions regarding program design and management issues. Second, helping the evaluation team to determine the target population and to clarify and update the project’s goals. Third, articulating the program theory, and then used that theory to develop measures for each program component. Fourth, assuring that stakeholders’ needs were addressed. Fifth, providing a basis to assess the significance of outcomes through ongoing assessments of the stakeholders’ needs. Sixth, targeting opportunities for use in project improvement.

Input Evaluation aided evaluators to improve their projects by: First, assessing those solution strategies/activities that were more effective in aiding the participants to meet the program needs. Second, selecting sources of support to assess how effectively were resources to meet the needs of the program. Third, assessing the strength of the project plans and resources. Fourth, improving the strategic planning and budgeting process. Fifth, assuring that that the project’s initial strategy was economically feasible for meeting the assessed needs of the different stakeholders.

Process Evaluation helped evaluators enhance their projects by: First, strengthening project implementation and design in areas of operational deficiency.
Second, implementing and refining the program design and procedures. Third, monitoring and maintaining records of the project’s process and costs.

Product Evaluation aided evaluators to improve their projects by: First, assessing the project’s success in reaching the appropriate stakeholders’ needs. Second, guiding decisions regarding termination, continuation, or modification of the project and to maintain a record of achievements. Third, examining the range, quality, duration, and importance of intended and unintended project’s effects on the stakeholders’ needs.

It should be noted that in terms of the similarities and differences in the outcomes found in both the BSC and CIPP evaluation models, the BSC is focused on a quantitative approach, while the CIPP tend to use qualitative methods in order to facilitate collaborative actions in the organization.

However, both models were used in these different organizations to escalate the visibility of quantified outputs and outcomes. The BSC provided an explicit linkage between strategic, operational, and financial objectives, and then communicated these linkages to managers and employee teams. Under the CIPP model, assessed needs of intended beneficiaries were identified and examined before an appropriate project strategy was developed and implemented. Outcomes were achieved through constant monitoring of both process and product. The CIPP provided support for a project’s improvement to these organizations in order to reach their short and long-term goals.
Organizations benefit from using the BSC or CIPP evaluation models when managers used them not only to communicate strategic direction and priorities, but also to provide a supporting decision-making framework.

Research Question # 4

What are the critical factors that seem to be associated with successful applications of the BSC and CIPP Model?

Throughout this dissertation, the term “critical factors” refers to the key issues associated with successful applications of both the BSC and CIPP models in order to aid managers and evaluators to identify a set of best practices of these models, when and how they are best applied, and to develop an awareness of how to improve these models.

Table 11.

Critical factors from BSC and CIPP evaluations

<table>
<thead>
<tr>
<th>BSC CASE # 1 SIEMENS AG</th>
<th>Critical Factors:</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼ Consistent executive support and involvement was critical in order to implement the BSC at Siemens. Senior management established a sense of urgency, by creating a guiding coalition, and developing a vision and strategy to guide behavior. After the BSC change process was launched, a revised governance system navigated the transition, followed by more permanent structural changes in the management system which affected resource allocation and compensation.</td>
<td></td>
</tr>
<tr>
<td>▼ A core solution, such as the Six Sigma model, was integrated with the BSC model, not only as a means to facilitate the transitions between analytic and operations tasks, but also to improve the impact of the BSC at Siemens.</td>
<td></td>
</tr>
<tr>
<td>▼ Selection of an appropriate organizational unit to implement the BSC model was critical. Siemens is has many different divisions and business units, constructing a corporate-level scorecard was a difficult first task. Thus, the initial scorecard process initiated in a strategic business unit (SBU). This SBU was selected as it was one that conducted activities across an entire value chain, including: innovation, operations, marketing, selling, and service.</td>
<td></td>
</tr>
<tr>
<td>▼ A few critical strategic objectives were identified during the development of the business strategy implementation phase at Siemens. Objectives are desired outcomes</td>
<td></td>
</tr>
<tr>
<td>▼ Process evaluation indicators were correlated with important outcome evaluation indicators. By linking the BSC to the Six Sigma program Siemens was able to focus on those “process strategic initiatives,” that were required to create value and to reach the targets.</td>
<td></td>
</tr>
<tr>
<td>▼ The BSC was a useful tool to focus efforts and communicate priorities. The BSC aided employees at Siemens not only to get strategic priorities clear, but also the Six Sigma model</td>
<td></td>
</tr>
</tbody>
</table>
### Table 11-Continued

<table>
<thead>
<tr>
<th>BSC CASE # 1</th>
<th>Siemens AG</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>was integrated as a core solution to identify root causes for problems and to weave a series of initiatives together.</strong></td>
<td></td>
</tr>
<tr>
<td>▼By implementing BSC at the business unit level; and then having senior management teams to serve as the connection between the executive team and employees, provided the means to ensure alignment as it relates to target agreement among different organizational levels.</td>
<td></td>
</tr>
<tr>
<td>▼The implementation protocol that Siemens used was the &quot;Four-Phase BSC Creation Process,&quot; including: business assessment, development of business strategy, measurement of the business strategy, and operations proved to be successful in accomplishing the strategic objectives.</td>
<td></td>
</tr>
<tr>
<td>▼Process improvement objectives were supported by strategically important key evaluation indicators.</td>
<td></td>
</tr>
<tr>
<td>▼Managers and employees at all organizational levels had the authority, responsibility and tools that were necessary to impact relevant measures. Leaders from each business unit were responsible for their BSC unit-level construction and implementation, which empowered them to have control over the process and their performance indicators.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BSC CASE # 2</th>
<th>Hilton</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Critical Factors:</strong></td>
<td></td>
</tr>
<tr>
<td>▼ A balanced scorecard is a system of linked objectives, measures, targets and initiatives which collectively describe the strategy of Hilton and how they were able to achieve their strategy.</td>
<td></td>
</tr>
<tr>
<td>▼BSC perspectives were customized to capture Hilton’s strategy. Each perspective reflected their unique organizational strategy. Hilton executives selected the following perspectives that drive value for the organization: operational effectiveness, revenue maximization, loyalty, brand standards, and learning and growth.</td>
<td></td>
</tr>
<tr>
<td>▼The BSC implementation process included: vision, value drivers, key evaluation indicators (property specific goals), and five C’s (constituencies).</td>
<td></td>
</tr>
<tr>
<td>▼Balanced scorecards were designed for all levels of employees in a hierarchical fashion, so that lower level measures link to higher ones. At Hilton, balanced scorecards were linked at the organizational, departmental, and individual level. The operations team was able to align everyone throughout the organization. Balanced scorecards were implemented at each hotel, and then cascaded to area and regional vice-presidents, and to the department manager and individual team member.</td>
<td></td>
</tr>
<tr>
<td>▼Balanced scorecards at every hotel within Hilton has the same “value driver focus,” but their key evaluation indicators were unique to their property, so that each hotel’s goals represented realistic and achievable performance levels.</td>
<td></td>
</tr>
<tr>
<td>▼Personal incentives were aligned, to ensure that employees at all levels were measured and compensated in ways that supported the strategy. For instance: balanced scorecards were linked to the hotel level manager and executive committee bonus and performance review to maintain focus and alignment throughout the organization.</td>
<td></td>
</tr>
<tr>
<td>▼The use of feedback mechanisms, such as the BSC software that was essential in order to manage the vast amount of information related to Hilton’s mission and vision, strategic goals, objectives, perspectives, measures, causal relationships, and initiatives. The use of a web-based BSC for measurement and information, provided managers with data that allow them to get to the root cause of problems, so that they were able to act on more accurate information.</td>
<td></td>
</tr>
</tbody>
</table>

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Successful implementation of the BSC Model was based on Kaplan & Norton’s five management principles to become a “Strategy-Focused Organization,” and included the following steps:

1. Mobilize change through executive leadership,
2. Translate strategy to operational terms,
3. Align the organization to the strategy,
4. Make strategy everyone’s job,
5. Make strategy a continual process.

BSC education and communication vehicles were critical factors to create strategic awareness and to measure its effectiveness throughout the organization. Mobil implemented, communicated, and cascaded its corporate scorecard that was designed by the executive, information technology, human resource teams, as a “pointer” into its 18 business units and 14 strategic partners.

Alignment of business units (corporate, business, departments, teams, and individuals) was critical to provide focus and to make sure that each part of the organization was lined up to best support the strategy. Alignment was critical to encourage employees empowerment and innovation because individual actions were directed at achieving high-level objectives.

Alignment required two steps:

1. Create awareness
2. Establish incentives.

First, the executive team communicated the high-level strategic objectives. Second, the executive team made sure that individuals and teams had local objectives and rewards that contributed to achieve the targets for high-level objectives.

Different levels of measures were used according to what each BSC owner were able to control. Each Strategic Business Unit’s (SBU) evaluation indicators were designed to stimulate corporate results in a cause-effect fashion, but not every evaluation indicator at the business unit level rolled up to an identical one on the corporate scorecard. This, helping to maintain the local relevance issues at each SBU.

Evaluation indicators were used to measure how Mobil was performing relative to its strategic objectives. Evaluation indicators were quantifiable performance statements. As such, they were:

1. Relevant to the objective and strategy
2. Placed in context of a target to be reached in an identified time frame
3. Capable of being trended
4. Owned by a designated person or group who had the ability to impact those evaluation indicators.

Mapping a strategy was an important way to evaluate and to make visually explicit Mobil’s perspectives, objectives, evaluation indicators, and the causal linkages between them. By organizing objectives in each defined perspective and mapping the strategic relationships among them, served as a way to evaluate objectives and to make sure they were consistent and comprehensive in delivering the strategy.

Education and communication were important factors to successfully implement the BSC at Mobil. A variety of communication vehicles were used to raise employee’s awareness of Mobil’s strategy.

Mobil linked compensation to scorecard-based outcomes, by using a bonus for up to 30% of base pay, applied to all salaried workers.

Mobil’s BSC was integrated with its annual planning, budgeting, and reporting cycle.
Table 11-Continued

<table>
<thead>
<tr>
<th>BSC</th>
<th>CASE # 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPS</td>
<td></td>
</tr>
</tbody>
</table>

**Critical Factors:**

▼ Evaluation indicators were consistent with research on the needs and priorities of UPS's customers, employees, shareholders, and other important stakeholders. UPS deployed a three-year, five phase program and focused everyone on the same "point of arrival" (corporate indicators) including: customer satisfaction, employee relations, competitive position, and time in transit.

▼ Establishment of mathematical correlations between key evaluation indicators such as customer or employee satisfaction and financial performance (including leading and lagging indicators). The BSC at UPS was used as a means to achieve a new measurement system in the organization that focused more on "process quality," and also helped to shift from a financial to a more strategic and predictive view of measuring performance customer or employee satisfaction and financial performance (including leading and lagging indicators). The BSC at UPS was used as a means to achieve a new measurement system in the organization that focused more on "process quality," and also helped to shift from a financial to a more strategic and predictive view of measuring performance.

▼ Balanced Scorecards at UPS included evaluation indicators that examine inputs, processes, outputs, and outcomes. UPS team were interested in developing a results-tracking rather than activity tracking system.

▼ UPS implemented the BSC by integrating it to their existing Total Quality Management system. After implementing the BSC, the quality and business plans that were formally separated, were integrated into one comprehensive business plan driven by the BSC.

▼ A few relevant and reasonable number of evaluation indicators were identified that outlined UPS strategy, and then the scorecards were strategically aligned at each organizational level. In order to create alignment among balanced scorecards at all levels, UPS executives sought evaluation indicators at each level which would directly contribute to the evaluation indicators at the next higher level in a cause-effect way. However, evaluation indicators were not the same at all levels.

▼ Through the quality performance reviews (QPR), evaluation indicators and goals were connected from the corporate level, district, business unit and finally to individual level. Thus, aligning balanced scorecards at the organizational, job, and individual level. In addition, the QPR served as a tool that functioned as an individual "business plan action."

▼ The use of indices (i.e., customer satisfaction index, safety index), helped management teams at UPS to simplify data and aid in the analysis. The BSC components were weighted differently according to the function and level of the person. Instead of having all quality performance review (QPR) items divided into equal 25% weightings for each BSC component at the corporate level, the role of the employee dictated the weightings for his or her own QPR document. Scorecards at the individual level achieved their "balance" among the four BSC components according to locally perceived role priorities. In all cases, the goal was to have a line-of-sight "impact visibility" from the front lines to the district level evaluation indicators for every employee.

<table>
<thead>
<tr>
<th>CIPP</th>
<th>CASE # 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIRIT OF</td>
<td>CONSUELO</td>
</tr>
</tbody>
</table>

**Critical Factors:**

▼ Successful implementation of the CIPP Model was based on Stufflebeam's four types of evaluation, and included the following components: context, input, process, and product evaluation.

▼ To assess project outcomes, product evaluation was divided into four parts: impact, effectiveness, sustainability, and transportability evaluation.

▼ The seven types of interrelated evaluation helped evaluators of the Consuelo Foundation to guide and assess its programs and projects. A context evaluation was conducted to establish the need for an intervention, and to provide the fundamental criteria for assessing the program success. Input evaluation was used to help assure that each project was well designed.
Table 11-Continued

<table>
<thead>
<tr>
<th>CIPP CASE # 1 SPIRIT OF CONSUELO</th>
</tr>
</thead>
<tbody>
<tr>
<td>to meet the target population’s assessed needs. Process evaluations examined the project staff’s execution of the project design and provided feedback for improvement. Evaluators employed impact evaluation to assess the project’s success in reaching the appropriate target audience. Evaluators also used effectiveness evaluation to examine the range, quality, duration, and importance of intended and unintended effects. Finally, evaluators applied sustainability and transportability evaluation to examine whether the project was resilient beyond the development period and attractive and feasible for use in other settings. Taken together, these seven kinds of evaluation provided a comprehensive basis for helping evaluators to conduct and sum up the merit and worth of this project.</td>
</tr>
</tbody>
</table>

▼Each type of evaluation was conducted to assess and to provide information to inform decision-making and accountability concerning a particular stage of the project development.

▼Formative evaluation was used to guide planning and to detect and correct problems. Thus, formative evaluation was useful for developing, managing, and improving ongoing efforts. Summative evaluation was used to assess the project’s success, and it examined how this project met the target groups’ needs, and how it achieved its objectives on improving the quality of life for children and families of the Waianae community. Formative evaluation was useful for developing, managing, and improving ongoing efforts. Summative evaluation was used to assess the project’s success, and it examined how this project met the target groups’ needs, and how it achieved its objectives on improving the quality of life for children and families of the Waianae community.

▼Evaluation indicators were consistent with the needs and priorities of stakeholders. Evaluation indicators were developed in the different types of evaluation (context, input, process, impact, effectiveness, sustainability, and transportability), to measure progress and impact of this project.

▼A project development cycle was used for project identification, planning, implementation, control, quality assurance, and dissemination. The use of this project cycle was an important way to evaluate and to make visually explicit the CIPP’s types of evaluation.

▼Evaluation questions were derived from the types of evaluation that guided this project, where context evaluation was used for goal-setting purposes, input evaluation was used for planning purposes, process evaluation was used to strengthen project implementation and design in areas of operational deficiency, impact evaluation was used for controlling purposes and to determine the extent to which this project reached the beneficiaries, effectiveness evaluation was used for assuring quality and to determine the project’s effects on the quality and life conditions of the Waianae community, and sustainability and transportability evaluation were used for institutionalizing and disseminating those successful aspects of this project that could be sustain and applied in other settings.

▼Short-term and long-term outcomes were achieved through enabling project drivers. For instance: Driver -Strategic Planning, Enabling Outcome-An effective strategic plan exists, is being implemented, and is conveyed clearly to others. Evaluation indicators included:
  1. Is the strategic plan linked to goals identified by the different stakeholder groups?
  2. Does the plan include a conceptual model or rationale for how the desired outcome would be achieved?

Short-Term Outcome:
The different stakeholder groups are involved and aligned with the Spirit of Consuelo program’s objectives and priorities.

▼Qualitative and quantitative methods for data collection were used for each component of the evaluation. A triangulation process was used, and each part of the evaluation was addressed by at least three different methods.
Table 11-Continued

<table>
<thead>
<tr>
<th>CIPP CASE # 1</th>
<th>SPIRIT OF CONSUELO</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼ Communication vehicles and feedback mechanisms such as annual reports including, environmental analysis evaluation report, project profile report, traveling observers reports, case studies, goal-free evaluation reports, and a final summative evaluation report, were used to encourage follow-through by the stakeholders of this project, and to increase the use of the evaluation results. Through these different communication vehicles the project evaluation team presented formative and summative feedback to the targeted groups.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CIPP CASE # 2</th>
<th>NASA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical Factors:</td>
<td></td>
</tr>
<tr>
<td>▼ This project used four types of the CIPP evaluation model (Stufflebeam, 1971), but focused primarily on the product evaluation component in order to gain additional insights into project outcomes. By focusing in the product evaluation component, the project evaluation team aimed at collecting descriptions and judgments of outcomes, to relate them to objectives, and to the context, input, and process evaluation information.</td>
<td></td>
</tr>
<tr>
<td>▼ The CIPP model was integrated with the “Formative-Summative” evaluation model (Scriven, 1967), to assess the impact and effectiveness of the Aerospace Education Services Program (AESP). Formative evaluation was used to improve the AESP program by answering the question of why AESP program results occurred and how can the program be improved? Summative evaluation was used to document program results, and focused on answering the question of what happened as a result of implementing the AESP program.</td>
<td></td>
</tr>
<tr>
<td>▼ The CIPP model was chosen as a means to obtain timely and credible information for purposes of decision-making. For instance: First, context evaluation was used for making decisions regarding program design and management issues (i.e., what problems or needs was the NASA-AESP attempting to address), the goals associated with meeting those clients' needs or using opportunities, and the objectives associated with solving problems (i.e., to what extent are the participants/program activities meeting the anticipated needs of the program) and to provide a basis for judging outcomes. Second, input evaluation was used for selecting sources of support (i.e., how effectively are resources (human and financial) being deployed to meet the needs of the program), solution strategies (i.e., what strategies/activities were more effective in aiding the participants to meet the program needs, and what barriers are being encountered), procedural designs; and providing a basis for judging implementation. Third, process evaluation was used for implementing and refining the program design and procedures (i.e., how, and to what extent, did the program and activities comply with state and local guidelines and frameworks, and contribute to support of state-level systemic education improvement efforts). Finally, product evaluation was used for making decisions regarding modification, or refocus of the AESP program education activities and services; and for presenting a clear record of effects intended and unintended (i.e., how have program activities impacted targeted K-12 students; to what extent were teacher workshops designed to include NASA curriculum materials; in what form, and how effective were program efforts to support informal education in institutions and organizations, teacher training programs, and other providers of related services).</td>
<td></td>
</tr>
<tr>
<td>▼ In order to avoid dependence on the validity of only one source of information, a triangulation process including multiple methods and sources was used to provide confirmatory measurement data to answer the evaluation questions included in this project. Qualitative and quantitative methods that were used in this project included: eight individual surveys (i.e., delphi survey was use in three different rounds), telephone surveys, document reviews, focus group meetings, direct observations, examination of existing databases. Sources of Information, included: NASA Head Quarters education administrators; OSU-AESP management team, NASA field center precollege officers; AESP specialists and administrative assistants; elementary and secondary students and; state-level science curriculum coordinators and leaders; other educators and local officials representing professional education organizations.</td>
<td></td>
</tr>
</tbody>
</table>
The NASA evaluation project, centered on 19 critical questions, and included four major phases over a three-year period. The four program areas included: support of systemic improvement, teacher preparation and enhancement that support systemic improvement, student support, and curriculum support and dissemination. These four phases were descriptive of the different types of evaluation work (i.e. context, input, process, and product), that occurred in this project.

Evaluation indicators were developed for the CIPP four-types of evaluation represented by the five major categories included in this project (i.e., program design and management, support of systemic improvement, teacher preparation programs and enhancement programs that support systemic reform, student support, and curriculum and dissemination). A representative sample of individual members from the different stakeholder groups were involved in determining these evaluation indicators of impact, thus empowering people to have control over the process and their own evaluation indicators.

Table 11 illustrates the analyses from the different case studies supporting several findings related to the critical factors associated with successful applications of the BSC and CIPP evaluation models.

**BSC Critical Factors**

The BSC’s model critical factors included: First, management support and involvement were critical in order to implement the BSC at these different organizations. Managers developed the vision and strategy to guide their employees’ behavior in the organization. Second, the implementation of the BSC model provided the mechanisms to integrate multiple management systems such as (TQM, Six Sigma) under a single decision-support and planning framework with a common strategic direction. Information from the BSC case studies pointed out that the BSC yield better results when it is linked with other evaluation models such as TQM or Six Sigma, as these include other methods to manage processes, and provide useful tools that are used particularly at an operational level. Third, the impact of a BSC depends to a great degree on the effective display of performance information. The use of web-enabled tools was critical to communicate information in these different...
organizations that have implemented the BSC model. In order to support and drive-goal achievement, these organizations moved towards a more knowledge-based orientation, and used these software tools to provide managers with timely and relevant data for decision-making processes. Therefore, tools used within the BSC (i.e., web portal), were mainly used for predicting outcomes.

Software systems provided fact-based information so that meaningful and accurate organizational strategies were formulated in order to make better decisions. Balanced scorecards are useful for managers as they pointed to critical problem areas that needed analysis, however managers have to perform that analysis and indicate what actions should be taken. In many cases, the BSC included evaluation indicators that were highly aggregated, and derived from multiple data points.

For example, at Siemens on-time delivery was below the targeted levels, corrective action was predicated by analysis of products, facilities, suppliers, or circumstances to determine who was responsible for those delays. These analytical capabilities were critical to improve performance at Siemens. However, the end goal of analysis is performance improvement through action, so there is a natural link between analysis and tools that prompt and facilitate action. However, it is important to point out that the use of BSC software tools is necessary but not sufficient. These tools does not replace the thinking needed from managers that is involved in implementing strategy and making decisions in organizations.

Fourth, impact of a BSC depends to a great degree on the alignment of evaluation indicators among the different organizational levels: organizational, departmental, and individual. Scorecards were designed for all levels of employees in
a hierarchical fashion so that lower level evaluation indicators link to higher level ones. Moreover, it is important to emphasize again that depending on the organization’s culture the BSC implementation process followed different approaches. Fifth, evaluation indicators focused on past, present, and future time perspectives; and the scorecard was “balanced” across the needs of customers, employees, and the shareholders of the organization. A few relevant and reasonable number of evaluation indicators were identified that outlined each organization’s strategy (usually between 12-16 evaluation indicators). Sixth, selection of an appropriate organizational unit to implement the BSC model was critical. The initial scorecard process usually initiated in a strategic business unit (SBU). The SBU selected by each organization was one that conducted activities across an entire value chain, including: innovation, operations, marketing, selling, and service. Each SBU’s evaluation indicators were designed to stimulate corporate results. Seventh, important evaluation indicators (i.e., those linked to scorecard factors) were given higher priority or weight in overall scorecard. The BSC perspectives (i.e., customer, financial, internal, and learning and growth), were weighted differently, according to the function and level of the person. Eight, the use of indices (i.e., safety index, customer index, employee index), helped management teams at these organizations to simplify data and aid in the analysis to simplify data and aid in analysis. Finally, establishment of correlations between key evaluation indicators such as customer or employee satisfaction and financial performance (leading and lagging evaluation indicators) were used in order to know what indicators were correlated with what
other, and how these linkages helped to improve business decisions.

CIPP Critical Factors

The CIPP’s model critical factors included: First, stakeholders support and involvement were critical in order to implement the CIPP at these different organizations. Stakeholders’ involvement in the identification and development of the project evaluation goals, objectives, and evaluation indicators that guided the evaluation’s implementation process was a key issue. Second, the implementation of the CIPP model provided the mechanisms to integrate different types of interrelated evaluation components including, context, input, process, impact, effectiveness, transportability, and sustainability. Third, the implementation of the CIPP model provided evaluators with a formative (prospective application of CIPP), and summative (retrospective use of CIPP) evaluation orientation for decision making improvement and accountability purposes. Formative evaluation’s purpose was to provide information to help develop and improve a program or project by supplying information for programming decisions and to assess and report on activities. The purpose of summative evaluation was to assess a program success. It examined whether the project met the stakeholders’ needs, objectives, and side effects. Fourth, evaluation questions were used to provide focus to the evaluation work. These critical evaluation questions (called value drivers in the BSC model) included the goals and objectives of the evaluation. Fifth, evaluation indicators in the CIPP model were developed based on the goals and objectives of the evaluation. Evaluators needed to identify and address the client’s purpose for the evaluation, which provided the key criteria for the evaluation indicators that were included in the evaluation. The CIPP
evaluation indicators on each of the different model’s evaluation components (context, input, process, and product) become the standards used to evaluate and communicate performance against expected results. Sixth, the CIPP Model used multiple qualitative and quantitative methods, and triangulation procedures to assess and interpret a multiplicity of information. These different methods applied in the CIPP Model were used in context, input, process, and product types of evaluation. The use of multiple methods for each type of evaluation provided crosschecks on findings. Seventh, the use of communication and feedback mechanisms were critical. Impact of a CIPP evaluation depends to a great degree on the use and dissemination of evaluation findings. Different means for disseminating findings include, oral reports, hearings, community forums, focus groups to examine and respond to findings, multiple reports targeted to specified audiences, and feedback workshops aimed at applying the findings. Eight, the CIPP model provided managers with flexibility in terms of which type of evaluation to employ, and what information to collect. The client’s needs and purpose for the evaluation, provided the criteria for the type(s) of evaluation components that were included in the evaluation. A summative evaluation required all four types of evaluation in order to fully describe the program and appropriately judge its quality. However, a formative evaluation assignment took up only the type(s) of evaluation needed to guide certain program decisions or answer pointed accountability questions.
CHAPTER V
DISCUSSION

Summary of Findings and Discussion

This study was concerned with reviewing the evaluation theory and practice of different evaluations models focused on improving strategic and operational management decision-making in organizations. Specifically this dissertation looked at the Balanced Scorecard (BSC), Kaplan & Norton, 1992; the Context, Input, Process, Product (CIPP) model, Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985; Total Quality Management (TQM), Deming 1920, 1982; Six Sigma, Welch, 1980, Anatomy of Performance (AOP), Rummler, 2002. To answer the questions of what were appropriate uses, interrelationships, and outcomes of evaluation models in practice the study integrated the BSC and CIPP Model. A thorough literature review provided an answer to research question 1 which focused on the differences and interrelationships among the BSC, CIPP, TQM, Six Sigma, and AOP evaluation models, and a review of BSC and CIPP’s case studies provided the basis for addressing research question 2 related to the similarities and differences related to actual implementation of BSC and CIPP Models in terms of its methods, including: evaluation components, evaluation indicators, data collected to support the evaluation, evaluation implementation protocol, and qualitative and quantitative analyses. Research question 3, was answered by determining the similarities and differences in outcomes of BSC and CIPP evaluation models. Research question 4, was answered by examining the critical factors that seem to be associated with successful applications of the BSC and CIPP Model.
In order to answer research questions 2-4, a multiple case study research method describing the context/applications, uses, methods, outcomes, of the BSC and CIPP models was conducted to answer the research questions included in this study. In addition, the Success Case Method (Brinkerhoff, 2003), was used as the evaluation tool to examine the different case studies in those organizations that have implemented the BSC and CIPP evaluation models.

The findings of this study are organized according to the three components proposed in this study: (1) Decision/accountability-improvement oriented evaluation models and strategic decision-making, (2) an overview of the BSC and CIPP evaluation model’s theory including the evaluation components, evaluation indicators, data collected to support the evaluation, evaluation implementation protocol, and qualitative and quantitative analyses and (3) merging Balanced Scorecard with the CIPP Model into a hybrid performance measurement evaluation model.

Decisions/Accountability-Improvement Oriented Evaluation Models and Strategic Decision-Making

The BSC, CIPP, TQM, Six Sigma, and AOP evaluation models were found to provide alternative ways of informing and improving both strategic and operational management decision making (Kaplan & Norton, 1992, 1996, 2004; Niven, 2003; Neely, 1998; Stufflebeam, 1971, Stufflebeam, Madaus, & Kellaghan, 2000; Rummler, 2002; Cokins, 2004; Evans & Lindsay, 1999; Pyzdek, 2003; Hayes, Austin, Houmanfar, & Clayton, 2001). These evaluation models differ in their theory and practice, including the model’s orientation, types of variables, information requirements, implementation protocols, and outcomes (Evans & Lindsay, 1999;
Klynveld, Peat, Marwick, Goerdeler, 2001; Pyzdek, 2003; Kaplan & Norton, 2004; Paladino, 2004; Cokins, 2004; Neely, 2004; Rummler, 2002). However, all of these evaluation models have a common purpose, they are all used to implement strategic performance evaluation that facilitates managers’ strategic decision-making, planning and control.

Not surprisingly, analysis of these evaluation models revealed that they have similar aspirations and concepts. They shared a number of characteristics such as, they were all measurement based, encourage a dialogue about strategic decision-making and performance improvement, strive to act as catalysts for change and action, and were also based on principles of on-going review, learning and feedback. Each of these evaluation models have been implemented in organizations for at least three reasons.

First, to maintain focus: evaluation models assisted managers by providing the elements or tools necessary to keep focus on those strategic objectives and higher-impact action steps and projects that will achieve those objectives, and by helping managers to define and use clear, concise performance indicators. These evaluation models provided managers with guidance about which direction the organization is going (i.e., by clarifying mission, vision, goals); which direction it should be headed (i.e., by defining strategic objectives, performance measures, and targets); and what it will require to get there (i.e., initiatives, management sponsorship and commitment, human capital, communication vehicles, technology infrastructure). By focusing on those critical areas, every employee within the organization were able to identify the sources of both an organization’s business problems and the opportunities that lead to
accomplish an organization’s strategic objectives. According to Neely (2004), “using strategic planning as a process to build a shared understanding around the critical few issues for an organization is key to successful implementation of the strategic plan implementation, and thus is key to organizational performance” (p. 170). Additionally, Kaplan & Norton (2004) noted that “The four-perspective model for describing an organization’s value-creating strategy provides a language that executive teams can use to discuss the direction and priorities of their enterprises. They can view their strategic measures, not as performance indicators in four independent perspectives, but as a series of cause-and-effect linkages among objectives in the four Balanced Scorecard perspectives” (p. 9).

Second, to reinforce communication of the strategic objectives: each of these evaluation models made use of elements and tools (i.e., strategy map, process mapping, and value driver tree tools), to aid managers in defining and communicating the strategic objectives and goals to employees in their organizations. Each model stressed the importance of selecting and defining those indicators and measures, each weighted to reflect its relative importance in relation with the strategic objectives and goals that were aligned with the organization’s mission and vision. A primary goal of these different evaluation models was to communicate the strategic vision to employees at all levels of the organization, and to empower employees to execute the organization’s strategy proactively. Different feedback mechanisms were used to integrate, distribute, and analyze the organizations’ information. Employees were then empowered to make decisions based on timely and relevant information. According to Cokins (2004), “employees and managers should be provided with the
tools to align their work with the strategy and to be recognized for their contribution to the organization’s success. A strategy-focused organization enables targeted feedback on strategic performance to specific employee teams, in order to effect continual strategy implementation” (p. 23). Moreover, Neely (2004) observed, “the definition of the performance creation process highlights the importance of creating alignment as a basic condition for an efficient use of resources and an effective trend towards the fulfillment of strategic intent (p. 78).

Third, to foster collaboration among employees for continuous improvement: these evaluation models were used to encourage a collective dialogue between managers and their employee teams regarding the implementation of the organization’s strategic objectives and performance indicators. “The specific meaning performance takes in an organization should be the result of extensive discussions between the various managers or decision makers of the organization. The goal of the discussions is to identify a coherent set of causal relationships and select a common set of indicators so that the coordination of all actors takes place and generates the value that, in the end, stakeholders define performance from their own point of view” (Neely, 2004, p. 78). Furthermore, Cokins (2004) stressed that, “the process cycle of managing strategy ends with collaboration. (The cycle never actually ends; it is a continuous iterative loop). By aligning various strategies among business units, the organization taps into the collective knowledge of its employees and unleashes each person’s potential” (p. 25).

These various evaluation models represented valuable tools for organizations to focus efforts and to communicate priorities, thus they helped managers to make
decisions regarding their programs and projects as it related to improving organizational performance and effectiveness. In brief, these evaluation models were used to aid managers align their own work as well as their employees' with the organization's strategy, mission, and vision, by focusing their efforts around those work activities that were critical to achieve long-term strategic objectives (Cokins, 2004).

BSC and CIPP Evaluation Models

In response to the second research question, related to the similarities and differences related to actual implementation of BSC and CIPP Model in terms of its methods including, evaluation components, evaluation indicators, data collected to support the evaluation, evaluation implementation protocol, and qualitative and quantitative analyses, the findings are as follows:

Evaluation Components

This study showed that both models have compatible evaluation components. For instance, evaluation components included in the BSC Model (Kaplan and Norton, 1992) included measures representing the four different BSC perspectives: customer, financial, internal, and innovation and learning, with some variations in the selection of these perspectives depending on the uniqueness of each organization's strategy. On the other hand, the CIPP model (Stufflebeam, 1971) employed as many as seven interrelated types of evaluation (context, input, process, impact, effectiveness, transportability, and sustainability) depending on the project evaluation's goals and objectives. Additionally, the CIPP model served formative, summative, or both purposes. Formative evaluation was used as a way to detect problems and weaknesses
in the project's components in order to revise them. Summative evaluation was employed as a process that concerned final evaluation to ask if the project or program met its goals. Figure 7, illustrate some parallels between BSC and CIPP model.

*Figure 7. Similarities between BSC and CIPP's evaluation components*

As depicted in Figure 7, BSC’s (internal, learning and innovation perspective) and CIPP’s (context, input, and process type of evaluation), represented the “enablers” areas which require people, resources, process, and activities investments in order to deliver the outcomes that an organization is trying to achieve. They are usually supporting activities within an organization that are not directly linked to the short and long-term outcomes. In addition, BSC’s (financial and customer perspective) and CIPP’s (product evaluation divided in impact, effectiveness, sustainability, and transportability types of evaluation), represented the “results” areas which expresses the needs of the stakeholders or customers, and identifies those components of value within the services and products offered by an organization.

**Evaluation Indicators**

In order to help managers and evaluators improve their strategic and operational decision-making processes, they require an accurate performance measurement system that provides them with evaluation indicators relevant and focused on the key strategic themes or objectives included in the strategic plan of the organization (in case of the BSC model) or in the evaluation project’s goals and objectives (in case of the CIPP model).

Evaluation indicators provide managers and evaluators with information to identify actions that should be taken in order to: accurately reflect the organization’s performance current situation, guide employees to take the right decisions in situations where action is required, and determine the effectiveness of those actions.
The organizations included in this study, used indicators as a means of developing a close loop feedback system that embodied situational analysis of information, corrective actions, and result evaluation.

The aim when creating and using indicators is to provide an objective view of an organization’s performance; thus, integrating qualitative and quantitative indicators is important as managers need to use quantitative evaluation indicators that will provide them with “hard quantitative” data to control the organization’s operations and financial areas, but also need the qualitative evaluation indicators that will provide them with a “through and in-depth analysis” data needed as a feedback mechanism for corrective actions and for creating value and employees’ behavioral alignment with the organization’s strategy.

Data Collected to Support BSC and CIPP Evaluations

The general context under which the BSC evaluation model was implemented in various organizations including the transportation, oil, hospitality, and transportation business/industry area. While the CIPP evaluation model was implemented at two different organizations including the community development and aeronautics and aerospace industry. The review of published BSC case studies showed that the initial scorecard process usually initiated in a strategic business unit (SBU). The SBU selected by each organization was one that conducted activities across an entire value chain, including: innovation, operations, marketing, selling, and service. Concerning the CIPP’s case studies, results showed that depending on the evaluation project’s goals and objectives the CIPP evaluation process started at the particular target audiences or groups that were identified by the client. Data collected
under the CIPP model included a comprehensive description and assessment of contextual information to determine the factors or conditions that influence the success of the organization’s goals, resources, processes, and products. Whereas the BSC focused on gathering customer, financial, operational, information to identify those strategic objectives and evaluation indicators that were aligned at the organizational, departmental, and individual level to support the organization’s strategy.

**Implementation Protocol**

The BSC model’s implementation protocol followed the five management principles of “Strategy-Focused Organization.” Alternatively, the CIPP model used four types of evaluations. The BSC model’s implementation protocol included the following principles: First; translating the organization’s strategy to operational terms (where evaluation indicators were defined for strategic objectives that comprised the organization’s BSC). Second; aligning the organization to the strategy (where all levels of the organization focused on strategic themes and priorities as defined by their strategy map and corresponding balanced scorecards for the corporate, business unit and support units). Third; motivation by empowering employees to implement the organization’s strategy (where individuals set personal work objectives which aligned with the organization’s BSC, and were rewarded with compensation and recognition for both individual and team accomplishment). Fourth; adaptation, by making strategy implementation a continual process (where managers linked the organization’s strategy to the budgeting process, management, and learning processes). Finally, mobilizing change through executive leadership (where managers
mobilized change in these organizations by demonstrating committed support to implement the BSC). Moreover, it is important to note that depending on the organization's culture the BSC implementation process followed either a top-down implementation (corporate BSC cascaded down to individual or employee level) or a decentralized implementation approach (middle-level managers and employee teams were responsible for defining the vision, value drivers, and performance indicators). The culture of the organization played an important role when determining which BSC implementation approach was to be followed; however; in all cases, a key issue for successful BSC implementation was employee involvement.

The CIPP model followed an implementation protocol based on four types of evaluation (context, input, process, product) (Stufflebeam, 1971, Stufflebeam & Shrinkfield, 1985) to assess and improve their evaluation projects including: context evaluation (that was used for goal-setting purposes and helped evaluators to determine the target population and to clarify and update the project's goals); input evaluation (that was used for planning purposes and aided evaluators to choose a program strategy); process evaluation (that was to help evaluators to strengthen project implementation and design in areas of operational deficiency) and product evaluation including in some cases an impact, effectiveness, sustainability, and transportability evaluation component (that was used for controlling, quality, and disseminating purposes. Additionally, the CIPP models employed also a formative and summative type of evaluation. Formative evaluation was used for project improvement, and summative evaluation was used to assess the project's success in terms of meeting the stakeholders' needs.
Qualitative and Quantitative Analyses

Some similarities and differences were found regarding the use of qualitative and quantitative tools that were employed in both models. One similarity that was found was related to the use of strategy maps (in BSC model), and project development cycle or logic maps in the CIPP model. These tools are used to help managers select and define strategic goals and objectives, and planned on the steps and projects that will achieve those objectives. For example, analyses from the different BSC case studies indicated that managers and employees used strategy maps (sometimes referred as process maps or value driver trees), as tools to develop focus in their organizations. By using these tools, managers and teams mapped cause-and-effect relationships for their organization’s strategy. On the other hand, the CIPP model used the project development cycle (sometimes referred as logic maps), as a tool used by evaluators for similar purposes that the BSC’s strategy maps are used for. The project development cycle was used by evaluators for: project’s identification and goal setting, planning, implementation, control and quality assurance.

A key point of difference between the BSC and CIPP’s tools that were used in these specific case studies, could be explained because of the particular quantitative focus of the BSC evaluation model. The different organizations that implemented the BSC model, generated many different scorecards at the corporate, divisional, departmental, and individual level. In order to point to critical areas that needed analysis, the balanced scorecards included measures that were highly aggregated and derived from multiple data points including details about strategies, objectives,
measures, and targets. Without the right software tools, this could be a complicated and confusing process.

The BSC and CIPP project management and quantitative and qualitative tools helped managers for operational, assessment/evaluation, and predictive planning purposes. Quantitative and qualitative analytical capabilities were critical to improve performance by providing information that managers can use to make decisions. The end goal of analysis is performance improvement through action, so there is a natural link between analysis and tools that prompt and facilitate action. Actionable and operational tools were used to facilitate action. Both the BSC and CIPP models used tools not only to motivate managers and employees, but also to empower them to make improvements.

The BSC evaluation model tended to use quantitative tools that were used to describe the logic of the linkage model, and then to translate it into the BSC’s key performance indicators, measures, targets, and initiatives. In contrast, the CIPP evaluation model employed mostly qualitative evaluation tools to gather data for each evaluation component.

A detailed summary of BSC and CIPP models (research question 2) is provided in Table 12.

Table 12.

Summary table of findings for BSC and CIPP model characteristics

<table>
<thead>
<tr>
<th>BSC Model Characteristics (General Findings)</th>
<th>CIPP Model Characteristics (General Findings)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context Dependent: tailored every time:</td>
<td>Context Dependent: tailored every time:</td>
</tr>
<tr>
<td>• An organization’s BSC was initiated by defining its own strategy and vision, and it</td>
<td>• The CIPP evaluation in an organization started by defining its own strategy and vision, and it</td>
</tr>
</tbody>
</table>
Table 12-Continued

<table>
<thead>
<tr>
<th>Prescriptive and Focused:</th>
<th>Descriptive and Comprehensive:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Every BSC was unique in its own objectives and measures for each organization.</td>
<td>• Every evaluation using the CIPP model was unique in its own objectives and measurement for each organization.</td>
</tr>
</tbody>
</table>

**Prescriptive and Focused:**

- The management team identified a unique and focused set of value drivers. Focus on the key evaluation indicators of success was essential. The focus of a BSC initiative is to shift from measurement of activities to measurement of results. Ultimately all of the organization's activities, resources, and initiatives should be aligned to the strategy.

**Descriptive and Comprehensive:**

- Comprehensive description and assessment of how context (goals), inputs (resources), processes (actions/activities), and products (intended and unintended outcomes) across an evaluation project were managed and deployed.

- There is no prioritization of one activity over another inherent in the CIPP model.

**Quantitative Fact-based:**

- The BSC is developed by having senior managers and employees make value judgments about how to reach the level of performance described in the vision and strategy of the organization.

- The BSC represented the insights, opinions, expertise and knowledge of the management team and employees with respect to the organization’s value drivers and quantitative evaluation indicators of success.

**Qualitative Fact-based:**

- The CIPP includes extensive qualitative information based on facts and data gathered within the organization. It is thoroughly documented from different sources of information (triangulation).

- Criteria and evaluation indicators used in the CIPP model’s case studies were not the same for any project. Both, criteria and indicators were defined by stakeholders, and were dependent on the program’s evaluation focus.

**Aspirational: “Should-be” perspective of the organization:**

- The BSC was built around a vision for what an organization wanted to achieve 2-5 years in the future.

- The BSC started at developing an organization’s vision, and then worked out back to the present to identify those performance gaps and the strategic roadmap.

- The BSC defined the required changes in measuring outcomes not only in the financial perspective, but also in the customer, internal, and learning and growth perspectives.

- The BSC assessed the current state in the quality of processes and activities of an organization by integrating other evaluation models such as TQM or Six Sigma that are used at the operational level of the organization.

**Current: “Present-Condition” perspective of the organization:**

- The Context and Input evaluation components in the CIPP model, described the current state of each evaluation project in terms of its current performance.

- The use of the CIPP model helped managers at organizations to identify relative strengths and weakness areas for improvement across the four types of evaluation (context, input, process, and product), and based on the previously stakeholders’ agreed upon criteria.

- CIPP outcomes do not pass judgment as to which activities need to be focused on a view of the future.

- The CIPP model encouraged continuous improvement in employees across the activities and operations on each project evaluation component.

157

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table 12-Continued

<table>
<thead>
<tr>
<th>Explicit Cause &amp; Effect Model:</th>
<th>Explicit Cause &amp; Effect Model:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The BSC model was set up in each of these organizations to drive explicit cause and effect objectives across the four perspectives, and managers made use of strategy map tools to represent each organization as a system of interlinked objectives.</td>
<td>• The CIPP model included a project development cycle used by evaluators for project’s identification and goal setting, planning, implementation, control and quality assurance.</td>
</tr>
<tr>
<td>• The cause and effect links were represented explicitly by using a logic model which included: resources, activities, outputs, short and long-term outcomes, and impact.</td>
<td>• External Variables were not systematically addressed:</td>
</tr>
<tr>
<td></td>
<td>• External variables such as the environment, and impact on society were not systematically addressed in the BSC model.</td>
</tr>
<tr>
<td></td>
<td>• External factors were included only if they were a key part of the strategy which needed to be tracked on a monthly basis.</td>
</tr>
<tr>
<td></td>
<td>• External factors were usually taken into account in setting targets for those measures included in the BSC of an organization (i.e., external benchmark, market research).</td>
</tr>
<tr>
<td>Source: Geraldina Villalobos, 2005. Doctoral Dissertation. Western Michigan University.</td>
<td>External Variables were systematically addressed:</td>
</tr>
<tr>
<td>Both evaluation models (BSC and CIPP) had a “context specific approach,” application in the different organizations. Both models were dependent and based on an organization’s competitive context, industry and economic challenges. The implementation either of the BSC or CIPP model, had to be customized for the organization’s unique circumstances.</td>
<td>• The Context evaluation component focus is on compiling and assessing relevant background information, and identifying needs and assets from contextual indicators to refine goals and objectives to better meet stakeholders’ needs. The input evaluation component focus is on profiling the program, reviewing pertinent research and development literature, and comparing the program’s plan with alternative strategies to help evaluators assure that initial strategies are defensible and feasible. These two evaluation components are an explicit part of the CIPP model, and therefore any evaluation project applying the CIPP model is going to provide feedback on these dimensions.</td>
</tr>
</tbody>
</table>

158

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
CIPP model provides a comprehensive assessment of an organization’s strengths and weaknesses, to aid managers in making decisions about those areas of improvement where they can focus their efforts. However, the CIPP model’s focus is independent of strategic priority, while the BSC model’s purpose is to help an organization to implement its strategy by identifying strategic objectives and translating those into evaluation indicators, which an organization needs to achieve in order to reach its vision. Attached to those strategic objectives is a set of actions and initiatives that the organizations needs to undertake to attain those objectives.

Outcomes of BSC and CIPP

Research question 3 focused on the similarities of outcomes between BSC and CIPP evaluation models, revealed that both models were implemented in organizations for the following three purposes: to maintain focus, assess and monitor performance, reinforce communication of the strategic objectives, and improve performance controls. One difference between the BSC and CIPP evaluation models was that the BSC model was found to be a strategic evaluation model for performance improvement which enabled an organization to articulate its strategy in a set of focused, strategic objectives and measures. The BSC evaluation model aimed at telling the story of an organization’s strategy in operational and measurable terms. Managers turned to the BSC model in an attempt to drive organizational focus and alignment of evaluation indicators from top to bottom throughout the organization.

On the other hand, evaluators used the CIPP model in an attempt to obtain credible information about their programs and services for purposes of decision making and accountability. The CIPP model focused on formative evaluation to

159
improve projects, as well as provide accountability records and summative evaluation reports that helped managers make informed decisions.

The BSC is focused on a quantitative approach, while the CIPP tended to use qualitative methods in order to facilitate collaborative actions in the organization. However, both models used a mixture of data types to escalate the visibility of quantified outputs and outcomes. The BSC provided an explicit linkage between strategic, operational, and financial objectives, and then communicated these linkages to managers and employee teams. Under the CIPP model, assessed needs of intended beneficiaries were identified and examined before an appropriate project strategy was developed and implemented. Intended and unintended outcomes were achieved through constant monitoring of both process and product. The CIPP provided support for a project's improvement to these organizations in order to reach their short and long-term goals.

Organizations benefit from using the BSC or CIPP evaluation models when managers used them not only to communicate strategic direction and priorities, but also to provide a supporting decision-making framework.

**Critical Factors of BSC and CIPP**

The greatest difference between BSC and CIPP models focused on the critical factors for successful implementation. A consistent theme found in the different BSC case studies was the importance of management support for successful implementation of this evaluation models. Another important factor for BSC implementation was related to the integration of other evaluation models such as, TQM or Six Sigma under a single decision-support and planning framework with a
common strategic direction. The impact of the BSC model depends to a great degree on the effective display of performance information; therefore, to support strategy evaluation, the use of software tools for reporting and analysis of the factors influencing the measures was a key issue. Alignment of evaluation indicators at all organizational levels was found to be critical, as it ensure that each part of the organization is lined up to best support the strategy. Evaluation indicators included in the BSC evaluation indicators focused on past, present, and future time perspectives. Founding a balance between lag indicators (which includes most financial measures) and lead evaluation indicators led to improved results in the organizations that implemented BSC. A few relevant and reasonable number of evaluation indicators were identified that outlined each organization’s strategy (usually between 12-16 evaluation indicators). Another critical factors was related to the implementation of the BSC evaluation model in an organizational unit that conducts activities across an entire value chain, including: innovation, operations, marketing, selling, and service. An important aspect that managers took into account when selecting the evaluation indicators included in the BSC was to weight these different evaluation indicators according to its relevance to the organization strategy, and the function and level of each person in the organization. Moreover, the use of indices was important as it aided managers and employee teams to simplify data.

Critical factors associated with a successful implementation of the CIPP model included stakeholders support and involvement. Stakeholders’ commitment to support the strategy played an important role in the identification and development of the project evaluation goals, objectives, and evaluation indicators. A key issue found
in the CIPP case studies was related with the CIPP model integration of different types of evaluation including, context, input, process, impact, effectiveness, transportability, and sustainability. Additionally, the implementation of the CIPP model provided evaluators with a formative (prospective application of CIPP), and summative (retrospective use of CIPP) evaluation orientation for decision-making purposes. Another important factor was the use of evaluation questions to provide focus to the evaluation work, by including the goals and objectives of the evaluation. Evaluation indicators were developed based on the goals and objectives of the evaluation. CIPP evaluation indicators were employed by evaluators as important feedback to evaluate and communicate performance against expected results. Additionally, the CIPP Model used multiple qualitative and quantitative methods, and triangulation procedures to assess and interpret information. The use of multiple methods for each type of evaluation provided crosschecks on findings. Another critical factor was the use of communication and feedback mechanisms, as impact of a CIPP evaluation depends to a great degree on the use and dissemination of evaluation findings. Finally, a critical factor when implementing the CIPP model was to assess the client’s needs before making a decision on which type of evaluation to employ and what information to collect under each evaluation component(s).

Evaluation Models and Strategic Decision-Making

A fundamental difference between the BSC and CIPP evaluation models, is that the BSC was design to communicate and assess strategic performance (in evaluation practice sometimes called “outcome-based evaluation), whereas the CIPP evaluation model can be regarded as a formal or disciplined approach to examine the value of a
program based not only on its outcomes but also on its context, inputs, processes and
procedures, and products (Worthen & Sanders, 1987). An evaluation makes use of a
systematic process of inquiry that includes developing the criteria or standards for
evaluation, the collection of relevant data and then making judgments about the
object of the evaluation by applying those previously developed standards in order to
determine quality.

Evaluation is made for the purpose of gathering information in order to make
rational decisions about changing elements of the program. In this interpretation of
evaluation the decision makers are fully intent on using data to alter the system, to
judge its value, and to change its direction if necessary.

For example, as part of assessing formative evaluation, the CIPP model,
seeks to establish how well an organization defines and manages the process of
strategic planning. It does this by determining whether the organization has a
formally established and appropriate process, which is reviewed regularly and is
systematically deployed at different levels. The BSC on the other hand, tests the
validity of the strategy and monitors the organization’s performance against its
delivery on a regular and frequent basis (i.e., monthly). The primary purpose of the
BSC is not to assess the quality of the strategic planning process itself, but to ensure
that the strategy gets implemented and to enable an organization to continuously learn
from its performance and adapt its strategy accordingly.

The BSC and CIPP Hybrid Evaluation Model

This study demonstrated that in the context of a conducting an evaluation
with the BSC or CIPP model, a hybrid model could be used (see Figure # 8) based on
a different methods, including: evaluation components, evaluation indicators, implementation protocols, qualitative and quantitative analyses, outcomes, and the critical factors associated with successful applications of the BSC and CIPP Model. Dynamics of the hybrid model capitalizes on BSC and CIPP evaluation model characteristics.

A hybrid model may have value due to similar philosophies regarding management, and may capitalize on different methods of measuring and managing an organization’s performance. Given this, a hybrid model will help evaluators and practitioners to understand some of these evaluation models’ distinctions in terms of: evaluation components, evaluation indicators, data collected to support the evaluation, implementation protocol, outcomes, and critical factors of each of these evaluation models. Figure 8, illustrates the hybrid model proposed in this study.

Figure 8. Hybrid model

![Diagram of Hybrid Model]


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Hybrid Model Characteristics. The hybrid model contained five components developed to articulate the expected similarities and differences between the BSC and CIPP models. The components of BSC and CIPP models pictured in Figure 8, were based on the findings of this dissertation. Table 13, provides an overview of the hybrid model’s components. Each of the components and general information on the distinctions in evaluation components, evaluation indicators, data collected to support the evaluation, implementation protocol, outcomes, and critical factors of each of these evaluation models elements is briefly discussed in this Table.

Table 13.

*Hybrid model’s components*

<table>
<thead>
<tr>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation Components:</td>
</tr>
<tr>
<td>The BSC’s (internal, learning and innovation evaluation components) and CIPP’s (context, input, and process evaluation components), are used as “enablers” elements which require people, resources, process, and activities investments in order to deliver the outcomes that an organization is trying to achieve. They include usually supporting activities within an organization that are not directly linked to the short and long-term outcomes. The BSC’s (financial and customer evaluation components) and CIPP’s (product evaluation divided in impact, effectiveness, sustainability, and transportability evaluation components), are employed as “results” elements which expresses the needs of the stakeholders or customers, and identifies those elements of value within the services and products offered by an organization.</td>
</tr>
</tbody>
</table>

| Evaluation Indicators: |
| The performance and evaluation indicators involved in each evaluation component in the BSC model (customer, financial, internal, and learning and growth), and the CIPP model (context, input, process, product), are used to improve an organization’s measurement system by helping managers and evaluators to develop relevant indicators that are focused on the critical factors included in the strategic plan of the organization or to improve an evaluation project. The BSC can be used for defining qualitative indicators and communicating strategy, and for measuring the effectiveness of the organization’s strategy implementation. The CIPP evaluation model can be used to collect information under the seven interrelated types of evaluation (context, input, process, impact, effectiveness, sustainability, and transportability), for defining qualitative indicators. Managers can use BSC and |
Table 13-Continued

Components

CIPP’s indicators as a means to develop close loop feedback systems to embody situational analysis of information, corrective actions, and result evaluation.

Implementation Protocol:
Depending on the performance evaluation context, the evaluator may employ a range of different performance evaluation’s tools and methods used in the BSC and CIPP evaluation models.

Outcomes:
This component represents the intended and unintended outcomes of both the BSC and CIPP evaluation models. This component provides also information on what are the goals. Thus, providing evaluators with an opportunity for to devise better alternatives and solutions to reach the desired outcomes.

Critical Factors:
The strengths and weaknesses provides evaluators with a critical review of what are some of the critical factors associated with successful applications of both the BSC and CIPP models, to help them identify a set of best practices of these models, when and how they are best applied, and develop an awareness of how to improve the models.


The hybrid model proposed seems to be sustained by those common themes found in these two evaluation models. Any kind of business (either in the industry sector or in the education area, private or public organizations) may use a BSC/CIPP hybrid model to improve their organization. The hybrid model provides an array of performance and evaluation tools that aid managers to implement strategic performance evaluation that facilitate managers’ strategic decision-making, planning and control.

Hybrid Model Implementation. First, within the context of different evaluation models for strategic decision-making, begin the evaluation of an
organization by seeking three distinct types of data (a) contextual, (b) implementation, and (c) outcomes data to support managers’ strategic decision making. Second, the CIPP model characteristics may be applied before the BSC model characteristics to gather information on the context and input evaluation components. Information under these components will add a deeper dimension to the BSC elements due to its comprehensive focus on compiling and assessing relevant background/environment information that includes pressures, expectations, constraints, and consequences. Stakeholders should be interviewed to determine congruencies and inconsistencies on strategic goals and objectives, and identifying needs and assets from contextual indicators that will aid managers assure that initial strategic themes are defensible and feasible. An organization that has implemented the hybrid model, will benefit by having a broadened understanding of its strengths and weaknesses at the context, input, process and product level. The CIPP interrelated evaluation types will provide guidance as to where an organization may need to improve significantly, where it performs adequately, and where it excels against the previously establish criteria.

Third, the hybrid model will shift to more closely resemble a BSC evaluation to provide focus and a clear plan of action to improve the organization’s measurement system. By using a strategy map, managers may communicate a unified view of the organization’s strategy to employees. The strategy map will aid managers to define the corporate direction and to align internal processes, strategic objectives, initiatives, evaluation indicators, and target scores. The BSC strategy map can be a valuable tool for providing strategic focus needed to prioritize action and to allocate resources. In
the above scenario, the BSC strategy map complements the CIPP process evaluation by providing a strategic prioritization tool for implementation of strategic objectives. By implementing the BSC, an organization’s can gain a through understanding of where to commit resources in order to impact those areas where the organization have been found to be particularly weak, but that support important strategic objectives.

Fourth, once the strategic objectives, initiatives, evaluation indicators, and target scores evaluation indicators have been defined using both the BSC and CIPP information including relevant background/environment information, strategic objectives, initiatives, evaluation indicators, and target scores, a web-based software tool can be used to display the organization’s system performance and evaluation and to communicate it to employees. These software tools will be used for predicting outcomes by providing managers with timely and relevant data for decision making processes.

Again, the BSC will complement the CIPP model by providing a strategic prioritization tool. By implementing the BSC, an organization’s can gain a through understanding of where to commit resources in order to impact those areas where the organization have been found to be particularly weak, but that support important strategic objectives.

On the other hand, the CIPP model will add a clear value on the BSC, by supporting those strategic objectives and measures, that were previously identified in the BSC. The CIPP model’s focus will be in providing a comprehensive approach to assess context, process, and products as central elements of the model. In addition, special attention will be given to the implementation process (formative evaluation),

168
including those activities or practices employed by managers in pursuit of achieving
the organization’s strategic goals. Therefore, the CIPP model can be very valuable
when addressing questions such as: How does the organization’s context contribute to
the organization’s outcomes?, How do organizational decisions and processes
employed by managers in their work, contribute to the organization outcomes? How
does productivity on intermediate outcomes contribute to the ultimate organization’s
strategic goals?

Each of these questions addresses the different elements included in the
hybrid model to achieve the organization’s outcomes. Within the hybrid model, data
would be collected along each of these elements, and a logic model can then be used
to illustrate these elements and corresponding links with the outcomes. In leveraging
the knowledge obtained from the hybrid model, managers can gain a depth of
understanding with respects to those challenges that they may face to deliver in their
attempt to reach the organization’s strategic objectives. For example, the hybrid
model could highlight particular areas of an organization’s weaknesses, which if there
are not addressed will make it difficult for an organization to reach its vision. This is
a valuable model to inform managers on how to fill the performance gap between its
current and desired performance. The CIPP model characteristics can provide
guidance on the type and level of efforts and investments required and the time frame
that it could take for the BSC objectives and measures to be implemented and fully
operational. The CIPP model could also act as an initial starting point to the BSC
implementation process.
It is important to emphasize again, that both the BSC and CIPP models add a useful dimension to the other by leveraging the knowledge and insights that each of these models bring about to managers for strategic decision-making purposes. The use of these models enrich managers’ knowledge by providing information about measuring and managing performance in an organization supported by an end to end analysis of an organization’s from strategy to operations and process improvement (Rummler, 2001).

This discussion illustrates how evaluators and practitioners in organizations may benefit from adopting these models and by adapting the BSC and CIPP to work together. However, it should be noted that given the distinct tools and approaches used in both models it is important that evaluators and practitioners should undergo an insightful process to determine why they are using these models, and where the differences are best addressed separately so that they can correctly manage their development and implementation, and most importantly help them to reach the organization’s goals.

Lastly, with the use of this hybrid model it is possible to identify a stronger fitting alternative model. These findings demonstrated that the proposed hybrid model can be used as a tool for conducting evaluation and performance improvement practices. Moreover, by examining the alternatives proposed in this hybrid model, this study aimed at providing some understanding of how long-term outcomes can be achieved in organizations. This study accomplished the goals of providing an alternative hybrid evaluation model, and provided a strong starting point for future research in this direction.
Limitations

There was one primary limitation of this study: sample characteristics. This limitation is presented as the foundation for the implications of this study, which address the fundamental issues that evaluators and practitioners should consider when implementing the BSC or CIPP evaluation models in organizations.

Within the context of this study, the units of analyses (the different organizations that have implemented either the BSC or CIPP model) were the different in the BSC and CIPP case studies. These units of analyses were appropriate in the context of the study given that answering the questions of what were the appropriate context/applications, uses, methods, products, strengths, weaknesses, and limitations of these evaluation models in practice.

In this study the sample was limited to the number of case studies that BSC and CIPP evaluators and practitioners have successfully implemented in different organizations including: Siemens AG in Germany (Power, Automation and Control, Transportation, Medical, Lightning and Information and Communications); Hilton Hotel Corporation (Hospitality/Services); Mobil North America marketing and refining (Oil corporation); United Parcel Services, UPS (Transportation Company); The Spirit of Consuelo: An Evaluation of Ke Aka Ho’Ona (Community Development and Self-Help House Construction); NASA-AESP (Aeronautics and Aerospace Industry), and that have documented their results and published them. These specific characteristics limited the sample of case studies included in this study.

The sample of case studies, however, were fundamentally representative of
those experiences of BSC and CIPP evaluators and practitioners, which provided a holistic portrayal substantial of those organizations that have implemented these evaluation models in order to learn from their experiences and results regarding the effectiveness of each of these models. Additionally, these different case studies were selected because they constitute exemplary illustrations of those organizations that have successfully implemented the BSC methodology.

Recommendation

It is recommended that the hybrid model be used by evaluators for their next strategic decision making evaluation. For researchers who are interested in adding to the body of knowledge on performance evaluation models, continued development of the hybrid model will facilitate better evaluation in the context of strategic decision-making.

Summary

The following summary points address what I view as the key lessons for integrating BSC and CIPP evaluation models:

1. The different evaluation models (i.e., BSC, CIPP, TQM, Six Sigma, and AOP), provided different approaches and tools that can be used by managers in their organizations for strategic decision making processes, that will be supported by fact-based data, and measured by outputs and outcomes.

2. Different benefits should be sought in using each of the models. Each model theory should incorporate context, process, and results, to comprehensively address why outcomes occur.
3. Above all, in using any evaluation model successfully, BSC, CIPP, TQM, Six Sigma, and AOP, there needs to be real and sustained management commitment. Without it, any evaluation model risks becoming an expensive, and short-lived exercise.

4. Development of evaluation indicators should be directly keyed to the organization’s strategy or project evaluation. A few relevant and reasonable number of evaluation indicators should be identified outlining each organization’s strategy.

5. Implement a full-organizational change strategy that integrates performance evaluation models with other core solutions. For example, merging BSC with other evaluation models such as CIPP, TQM, Six Sigma, and AOP.

6. Integrate evaluation models with other organizational systems such as performance appraisal, cost-management.

7. The implementation and evaluation of an organization’s strategy poses challenging problems, that require new performance and evaluation models and tools to keep up with the rapid and increasing complexity of an organization’s system. Therefore, the use of these evaluation models may provide managers with the necessary tools that can be used for strategic decision-making processes.
REFERENCES


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


St. Lawrence, D., & Stinnett, B. (July, 1994). Powerful planning with simple techniques. *Quality Progress*, 27 (7), 57-64.


APPENDIX A

Source Information for BSC and CIPP Model Case Studies
BSC Case Study 1. Siemens AG in Germany


Industry: Power, Automation and Control, Transportation, Medical, Lightning and Information and Communications.

BSC Case Study 2. Hilton Hotel Corporation.


Industry: Hospitality/Services

BSC Case Study 3. Mobil North America Marketing and refining (Mobil NAM&R).


Industry: Oil Corporation

BSC Case Study 4. United Parcel Services (UPS)


Industry: Transportation Company
CIPP Case Study 1. The Spirit of Consuelo: An Evaluation of KE AKA HO’ONA

Source: Stufflebeam, Daniel; Gullikson, Arlen; and Wingate, Lori. (2002). The Spirit of Consuelo: An Evaluation of Ke Aka Ho’ona. Western Michigan University and The Evaluation Center, Kalamazoo, Michigan 1-154

Industry: Community Development and Self-Help House Construction.

CIPP Case Study 2. NASA-AESP

Source: Horn, Jerry. G and McKinley, Kenneth H. (October, 2004). Evaluation of the National Aeronautics and Space Administration Aerospace Education Services Program (NASA-AESP). The Evaluation Center and Western Michigan University, Kalamazoo, Michigan 1-125

Industry: Aeronautics and Aerospace Industry.