Vladimir Mandić

MEASUREMENT-BASED VALUE ALIGNMENT AND REASONING ABOUT ORGANIZATIONAL GOALS AND STRATEGIES

STUDIES WITH THE ICT INDUSTRY
VLADIMIR MANDIĆ

MEASUREMENT-BASED VALUE ALIGNMENT AND REASONING ABOUT ORGANIZATIONAL GOALS AND STRATEGIES
Studies with the ICT industry

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Technology and Natural Sciences of the University of Oulu for public defence in Wetteri-sali (Auditorium IT115), Linnanmaa, on 14 September 2012, at 12 noon

UNIVERSITY OF OULU, OULU 2012
Software and ICT companies fail due less to technological factors than socioorganizational factors, among which the most common is unrealistic goals and objectives. The socioorganizational factors are present in companies and organizations of all sizes. Organizations need to know whether their goals and strategies are working and whether the strategies are effectively providing a reasonable return on investment for the effort that is being applied, i.e. to understand how valuable they are for the organization. In addition, organizations need to recognize what the risks are in achieving those goals and evaluate their probability. And, when goals are set properly then they have directive and energizing functions, they tend to utilize available resources better and serve as activators of cognitive processes and knowledge sharing cycles.

In this dissertation the GQM+Strategies approach was used. The approach was designed to help organizations to align goal-driven measurement schemes, i.e. GQM, with organizational goal hierarchy. The extensions developed in this dissertation (i.e., the solution proposed here) evolve the GQM+Strategies approach by providing an organization with capabilities to: (1) apply the work of value-based software engineering to directly address the return on investment (ROI) of their goals and strategies via evaluation of the costs and benefits of the goals and strategies chosen, (2) calculate a set of earned value metrics that allows organizations to effectively monitor the implementation of the organizational goals and strategies, (3) identify the risks associated with not achieving various sub-goals in a grid by analyzing goal risk exposures and acceptable risk levels, and (4) assess the threats of risky goals on other goals in the grid via evaluation of goal dependencies using the formal goal-strategies-goals models.

The GQM+Strategies approach was piloted in four different organizations involving more than 60 participants. The feedback from the participants was used to identify the research questions posed in this dissertation. The research approach (solution development process) adopted the design science framework and utilized analytical and empirical paradigms in different phases of the solution development. The analytical paradigms were used for solution development and evaluation, while empirical paradigms where used for evaluating certain aspects of the solution.

Keywords: earned value, goal-driven measurement, GQM, GQM+Strategies, metrics, organizational reasoning, strategic alignment, Value-Based Software Engineering, VBSE
Tiivistelmä


Tässä väitöskirjassa käytettiin GQM+Strategies lähestymistapaa. Tämä lähestymistapa on suunniteltu auttamaan organisointiota linjaaan tavoitelähtöisiä mittauksenmenetelmiä, kuten GQM, organisaation tavoittehierarkian suhteen. Tässä väitöskirjassa esitetty laajennus kehitettiin GQM+Strategies lähestymistapaa tarjoamalla organisaatioille kyvykkyyksiä: (1) soveltaa arvoperustainen ohjelmistotuotannon menetelmiä tavoitteiden ja strategioiden investoinnin suotujen tarkasteluihin arvioimalla valittujen tavoitteiden ja strategioiden kustannuksia ja hyötyjä, (2) laskea joukkou suoautetun arvon mittareita jotka mahdollistavat organisaatioille toteutettujen organisaatiolisaiden ja strategioiden tehokkaan seurannan, (3) identifioi riskieitä jotka liittyvät eri tavoitteiden saavuttamattomuuteen analysoidakseen arvioida riskejä jotka hyväksyvät riskitäkkätä riskitasojen ja (4) arvioida riskieitä sisältävien tavoitteiden uhkaa suhteessa muihin tavoitteisiin arvioimalla tavoitteiden riippuvuuksia käyttäen organisaatioilla tarjoamaa asemaa strategiat-tavoite-malleja.

GQM+Strategies lähestymistapaa pilottiin neljässä organisaatioissa yhteensä yli 60 osallistujan kanssa. Tutkimuksen osallistujilta saatuja palautetta käytettiin väitöskirjan tutkimuskysymysten identifioinniin. Tutkimuksen lähestymistavassa (ratkaisun kehittämisprosessi) sovellettiin design science -kohdokaa ja ratkaisun kehittämisen eri vaiheissa hyödynnettä analyysitisiä ja empiiristä paradigmoja. Analyysitisiä paradigmoja käytettiin ratkaisun kehittämisessä ja arviointinassa ja empiirisiä paradigmoja ratkaisujen määrittävien aspektien arviointissa.

Asiasonat: ansaittu arvo, arvoperustainen ohjelmistotuotanto, GQM, GQM+Strategies, metriikka, organisaatioalainen päättely, strateginen suuntaaminen, tavoitelähtöinen mittaaminen, VBSE
To my mom and dad who made all this possible
Preface

Writing a dissertation is a life experience that goes beyond the 310-page monograph. During my PhD journey, I met new people and colleagues, and forged new friendships. The journey started at University of Oulu, Department of Information Processing Science in early 2008, and lasted for four and a half years. Now, I evoke memories and express my gratitude.

First, I would like to thank Prof. Markku Oivo and Dr. Jouni Markkula, who gave me an opportunity to work as a researcher on the VASPO project. That project had a significant impact on my dissertation work. Through that project, I shaped the topic of my dissertation, I received support from my advisors Markku and Jouni, and I had the chance to collaborate with Prof. Vic Basili, who became my external co-advisor. I appreciate the support and freedom that my advisors gave me to explore my own ideas. Especially, I have to thank Vic for showing me the beauty and power of a simple academic thought, and for guiding me with paternal care through my PhD journey. There were moments when I really needed such care; therefore, once again, thanks, Vic.

I could not do my research without the energizing and inspiring working environment at the Department of Information Processing Science, and all my colleagues in the department. The best research group to work with in the department is the M-group. Pilar, Burak, Harri, Mika, Pasi, Markku, Samuli, Jouni M., Jouni S., Kari, Sanja, Markus, Jarkko, and Lasse are colleagues from the M-group with whom I had the pleasure of collaborating, and I hope to continue our collaboration in the future. This list would not be complete without mentioning my officemates. The first couple of years, I shared my working space with Lech, Mikko, and Tero. After we moved into the new building, Burak became my new officemate, and he got a better desk position, because Pilar flipped a coin in his favor. I will always remember freezing coffee breaks with Burak and cool academic discussions. I mean literally "cool"; outdoors it was down to -20 °C.

Moving to a new country for an extended period has consequences for one’s social life. The people who helped me the most to better understand Finnish culture and the social habits of Finns are Sanja and Harri. In addition, Dan Bendas made my process of moving to the North less painful with very precise suggestions and personal help. He was someone I could rely on when I needed something. The best way to adopt Finnish culture is by years of practicing Finnish sauna rituals. Therefore, I thank all
my colleagues who patiently taught me the proper ways of taking part in and enjoying saunas. The northern winter nights would be much darker without our social dinner gatherings. I will definitely miss that, as well as the company of Pilar, Burak, Harri, and Nebojsa.

During my PhD studies, I changed my accommodation several times in Finland and once in the United States. I would like to thank my flatmates Alberto, Nebojsa, Leo, and Wu who made living abroad feel like home.

Without financial support, I would not have been able to finish my dissertation. Foremost, I have to acknowledge the SoSE (Graduate School of Software Engineering) funded position. And the research projects VASPO and CloudSW, both Tekes-funded projects, for the opportunity to work with Finnish industry.

From September 2011 until March 2012, I worked as visiting scholar at the University of Maryland, in the United States. There, I had the pleasure of collaborating closely with Prof. Vic Basili, Dr. Carolyn Seamen, and Dr. Madeline Diep. The trip to the States would not have been possible without financial support from the University of Oulu and Tönninging säätiö grants.

I would like to thank Dr. Jyrki Kontio and Dr. Forrest Shull, my dissertation manuscript reviewers, for doing extensive reviews and providing suggestions that not only improved the manuscript but also will affect my future research on this topic. Last but not least, I have to thank my friend Dr. Nebojsa Gvozdenovic, a mathematician who helped me polish the formulas and mathematical notation in the dissertation.

Living abroad is a new and exciting experience, but at the same time made me realize the importance of family and friends back home. I would like to thank them all for having trust in me that I am doing the right thing. Finally, at the end of my PhD journey, I can say that I did the right thing.

In Novi Sad, Serbia, 16.08.2012.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>GQM</td>
<td>Goal Question Metric</td>
</tr>
<tr>
<td>VBSE</td>
<td>Value-Based Software Engineering</td>
</tr>
<tr>
<td>QIP</td>
<td>Quality Improvement Paradigm</td>
</tr>
<tr>
<td>EF</td>
<td>Experience Factory</td>
</tr>
<tr>
<td>ESE</td>
<td>Empirical Software Engineering</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technologies</td>
</tr>
<tr>
<td>NATO</td>
<td>North Atlantic Treaty Organization</td>
</tr>
<tr>
<td>PDCA</td>
<td>Plan-Do-Check-Act</td>
</tr>
<tr>
<td>ACMR</td>
<td>Alignment, Communication, Measurement, and Reasoning</td>
</tr>
<tr>
<td>TSP</td>
<td>Team Software Process</td>
</tr>
<tr>
<td>EVM</td>
<td>Earned Value Management</td>
</tr>
<tr>
<td>EVA</td>
<td>Earned Value Analysis</td>
</tr>
<tr>
<td>CMMI</td>
<td>Capability Maturity Model Integration</td>
</tr>
<tr>
<td>CobIT</td>
<td>Control Objectives for Information and Related Technology</td>
</tr>
<tr>
<td>ISO</td>
<td>International Standard Organization</td>
</tr>
<tr>
<td>BSC</td>
<td>Balanced Scorecard</td>
</tr>
<tr>
<td>SWEBOK</td>
<td>Software Engineering Body of Knowledge</td>
</tr>
<tr>
<td>IEEE</td>
<td>Institute of Electrical and Electronics Engineers</td>
</tr>
<tr>
<td>ACM</td>
<td>Association for Computing Machinery</td>
</tr>
<tr>
<td>SEI</td>
<td>Software Engineering Institute</td>
</tr>
<tr>
<td>PSM</td>
<td>Practical Software Measurement</td>
</tr>
<tr>
<td>ITIL</td>
<td>Information Technology Infrastructure Library</td>
</tr>
<tr>
<td>SEL</td>
<td>Software Engineering Laboratory</td>
</tr>
<tr>
<td>BTOPP</td>
<td>Business, Technology, Organization, Process, and People</td>
</tr>
<tr>
<td>C/SCSC</td>
<td>Cost/Schedule Control Systems Criteria</td>
</tr>
<tr>
<td>BCWS</td>
<td>Budgeted Cost of Work Schedule</td>
</tr>
<tr>
<td>ACWP</td>
<td>Actual Cost of Work Performed</td>
</tr>
<tr>
<td>BCWP</td>
<td>Budgeted Cost of Work Performed</td>
</tr>
<tr>
<td>CPI</td>
<td>Cost Performance Index</td>
</tr>
<tr>
<td>SPI</td>
<td>Schedule Performance Index</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>Research and Development</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>ARE</td>
<td>Acceptable Risk Exposure</td>
</tr>
<tr>
<td>RE</td>
<td>Risk Exposure</td>
</tr>
<tr>
<td>PBRS</td>
<td>Planned Benefits Realization Schedule</td>
</tr>
<tr>
<td>ABRM</td>
<td>Actual Benefits Realization Materialization</td>
</tr>
<tr>
<td>PBRM</td>
<td>Planned Benefits Realization Materialization</td>
</tr>
<tr>
<td>BPI</td>
<td>Benefits Performance Index</td>
</tr>
<tr>
<td>PDT</td>
<td>Probability Distribution Table</td>
</tr>
<tr>
<td>DAG</td>
<td>Directed Acyclic Graph</td>
</tr>
</tbody>
</table>
## Contents

Abstract  
Tiivistelmä  
Preface  
Abbreviations  
Contents

### 1 Introduction

1.1 Motivation and Research Opportunities  
1.1.1 Value-Based Software Engineering  
1.1.2 Goal-Driven Measurement Alignment  
1.2 Overview of the Research Questions and Research Approach  
1.3 Terminology  
1.4 Notes on the Dissertation Structure

### I Related Work and Research Approach

2 Related Work

2.1 Literature Sources and Review Methodology  
2.1.1 Reference-Based Literature Search Protocol  
2.1.2 Systematic Literature Review Protocol  
2.2 Goal-Driven Alignment and Measurement  
2.2.1 Review of Goal-Driven Alignment Approaches  
2.2.2 Review of Goal-Driven Measurement Approaches  
2.2.3 The Gap between Measurement and Organizational Goals  
2.3 GQM+ Strategies Approach  
2.3.1 The Meta-Model  
2.3.2 The GQM+ Strategies Process: The Grid Derivation Process  
2.4 Value-Based Software Engineering  
2.4.1 The "4+1" Theory of the Value-Based Software Engineering  
2.4.2 Theory W  
2.4.3 Value-Based Software Metrics  
2.4.4 Value-Based Risk Management  
2.5 Design Science  
2.5.1 Work of Herbert A. Simon
2.5.2 Hevner et al.'s Research Framework ........................................... 83
2.6 Empirical Software Engineering Methods ....................................... 86
  2.6.1 Participant Observation ............................................................. 88
  2.6.2 Interviews .................................................................................. 88
  2.6.3 Surveying ................................................................................... 89
  2.6.4 Coding Techniques ..................................................................... 89
3 Research Approach ........................................................................ 91
  3.1 Research Setting: the Role of Theory and Practice ......................... 93
  3.2 Research Process ........................................................................... 95
    3.2.1 Phase I: Defining the Research Problem .................................. 96
    3.2.2 Phase II: Developing the Solution ............................................ 97
    3.2.3 Contributions to Theory and Solution Usefulness .................... 99
  3.3 Used Research Methods ............................................................... 101
II Defining the Research Problem ..................................................... 103
4 The Pre-Study: Empirical Investigations of the GQM+Strategies approach ................................................................. 105
  4.1 The Pre-Study Design .................................................................. 106
    4.1.1 Technology Transfer ................................................................. 106
    4.1.2 Data Collection ......................................................................... 108
    4.1.3 Data Analysis ........................................................................... 109
  4.2 GQM+Strategies Industry Pilots ...................................................... 115
    4.2.1 Industry Case 1: Telco#2 ......................................................... 116
    4.2.2 Industry Case 2: ITDep#1 ......................................................... 117
  4.3 Data Collection ............................................................................... 118
  4.4 Data Analysis and Findings ............................................................ 118
    4.4.1 Finding #1: disjoint goals and strategies are posing a real challenge for the organizational alignment and transparency........ 119
    4.4.2 Finding #2: the GQM+Strategies grid development process helps identify and resolve the issue of organizational misalignment ................................................................. 120
    4.4.3 Finding #3: organizational goals whose values are not expressed by financial indicators tend to become less important than the goals whose values are quantified by financial indicators ................................................................. 121
4.4.4 Finding #4: the assumption elements are useful in coping with
the complexity and uncertainty of the business environment. . . . . .122

4.4.5 Finding #5: better mechanisms for incorporating the risk
information into the GQM*Strategies grids are needed. . . . . .123

4.4.6 Finding #6: a mechanism for quantifying the effectiveness of
strategies is needed. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .123

4.5 Reflections and Observations . . . . . . . . . . . . . . . . . . . . . . . . . . . . .124

4.6 Research Questions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .125

4.6.1 RQ1: Enunciating Value Propositions . . . . . . . . . . . . . . . . . . . . . . . . .126

4.6.2 RQ2: Monitoring Organizational Earned-Value . . . . . . . . . . . . . . . .126

4.6.3 RQ3: Analyzing Value Proposition Dependences . . . . . . . . . . . . . . . .127

5 Research Problem and Design Hypotheses 129

5.1 Research Problem . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .129

5.2 Design Hypotheses and Activities . . . . . . . . . . . . . . . . . . . . . . . . . . . . .131

5.2.1 Design Hypotheses . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .132

5.2.2 Activity 1: adding a value-based component to the
GQM*Strategies approach . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .135

5.2.3 Activity 2: building a model for monitoring cost, benefits, and
risk of the grid elements . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .135

5.2.4 Activity 3: constructing the goals–strategies dependences
models of the GQM*Strategies grids . . . . . . . . . . . . . . . . . . . . . .135

III Developing the Solution 137

6 Organizational Value Alignment and Measurement with the
GQM*Strategies 139

6.1 Problem Definition . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .140

6.1.1 Subproblem 1: Enunciating Organizational Value Proposition . . . .141

6.1.2 Subproblem 2: Integrate Cost and Benefits Estimates of Value
Propositions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .144

6.1.3 Subproblem 3: Value Realization Risk . . . . . . . . . . . . . . . . . . . . . . . . . .144

6.2 Solution Part I: Enunciating Organizational Value Proposition . . . . . . .144

6.2.1 Defining the Hierarchy of Goals with the Grid Derivation
Process . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .145

6.2.2 Formalizing the Value Propositions with the GQM*Strategies
Goal Template . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .145
6.3 Solution Part II: Integrate Cost and Benefits Estimates of Value

6.3.1 The Value Dimensions of the GQM* Strategies Goal

6.3.2 Integrating Cost Estimates

6.3.3 Taking Benefits into Account

6.4 Solution Part III: Value Realization Risk

6.4.1 Risk-Related Dimensions of the GQM* Strategies Goal

6.4.2 Incorporating Risk into the GQM* Strategies Interpretation Model

6.5 Evaluation of the Approach

6.5.1 Analytical Evaluation: GQM* Strategies as Notation for the Value-Based Software Engineering

6.5.2 Findings

7 Organizational Earned-Value Metrics

7.1 Problem Definition

7.1.1 Subproblem 1: Tracking Actual Costs and Benefits

7.1.2 Subproblem 2: Organizational Earned-Value Metrics and Indices

7.2 Solution Part I: Tracking Actual Costs and Benefits

7.2.1 The Cost–Benefit GQM Graph

7.2.2 The Notation of the Goal’s Cost and Benefit Metrics

7.3 Solution Part II: Organizational Earned-Value Metrics and Indices

7.3.1 Basic Organizational Earned-Value Metrics

7.3.2 Cost and Benefits Performance Indices

7.3.3 Interpreting Organizational Earned-Value Metrics: Analyzing Effectiveness of Strategies

7.4 The Compliance of the Organizational Earned Value Analysis with Value Based Monitoring and Control

8 Constructing Organizational Causal Models with the GQM* Strategies

8.1 Problem Definition

8.1.1 Subproblem 1: Identifying the right question

8.1.2 Subproblem 2: Calculate missing PDT values

8.2 Solution Part I: Identifying the Right Question

8.2.1 Modeling Events with the GQM* Strategies Concepts
1 Introduction

What we have to learn to do, we learn by doing.
by Aristotle

Nowadays, the software industry is doing a better job of producing high-quality products within budget and schedule than ten years ago. Still, according to Standish Group\(^1\), the situation is far from satisfactory. The *Chaos Reports*\(^2\) for the period of 1994–2009 show an improvement in project success figures, from 1994, when 16% of the projects were categorized as successful, meaning they were completed on time, on budget and met user requirements, up to 35% and 32% success rates in 2006 and 2009. Although there are serious studies that express concerns regarding the precision of the Chaos reports [73], it is evident that success performance in the software industry fluctuates. The Standish Group reports are observational reports, which monitor trends in industries, but they do not provide deeper insights into why those trends occur.

Ewusi-Mensah published the *Software Development Failures* book from an extensive MIT study on failed projects in the period from the late 1980s until the 2000s [74]. The study included a large number of multiple sources from surveys conducted by the author to the meta-analysis of published results. The analysis framework was based on three categories of critical factors: socioorganizational, sociotechnological, and economical [74]. The study showed that socioorganizational factors had the strongest impact on project failure. The collected data (evidence) support a conclusion that pure technological factors are not the main problem at all, but rather non-technological factors are. Interestingly, those factors are equally present in large corporations as well as in small and medium-sized companies [74]. For example, one of the socioorganizational failure factors is *unrealistic project goals and objectives* [74, p.51], which in turn impacts further downstream project decisions. Similar findings about the software project failures are reported throughout literature, e.g. [6, 24, 141, 155, 200].

The decision-making process in organizations involves developing an effective set of operational goals and defining strategies to implement them [16, 22]. Organizations need to know whether these goals and strategies are working and whether the strategies

\(^1\)www.standishgroup.com
http://www.sdtimes.com/content/article.aspx?ArticleID=36247
are effectively providing a reasonable return on investment for the effort that is being applied [22, 94, 182, 183, 187]. In addition, they need to recognize what the risks are in achieving those goals and evaluate their probability [22]. However, this process is far more complicated than a simple analysis of options to satisfy an expected value function (e.g., utility function); it involves the identification and valuation of numerous variables, from individual knowledge, experience, and affinity variables to group and organizational context variables, many of which are hard to quantify [22]. Therefore, organizations are often unable to recognize the return on investment (ROI) when they make decisions, and once they make decisions, they lack the means to effectively assess and manage the risks of the activities resulting from the decisions.

Furthermore, organizations need to communicate those goals and strategies throughout the organization, translating higher-level goals and strategies into synchronized goals and strategies at every level of the organization [182]. All parts of the organization need to move in the same strategic direction [120]. To ensure that their goals and strategies are aligned and successful, they should be able to measure and make visible their effects and control their implementations, i.e., recognize the achievement of their goals and assess the effectiveness of their strategies, and make improvements where necessary.

This is even harder for creative industries [102, 158], like software companies, where there may be different value propositions for different stakeholders at different levels of the organization [37, 182, 187]. Therefore, each organization needs to synchronize where necessary, but focus on local issues to balance individual needs [22, 40, 182]. Software companies rely on the knowledge, experience, and creativity of individuals and groups to produce end-products or services. Such a setting is aptly described by Nonaka & Takeuchi [158] as a knowledge-creating company, or what can be referred to as a creative industry. In order to be successful, companies in the creative industry need to [182]: provide and communicate a holistic view of the business and operations to creative people across the organization, manage creative people, and utilize metric programs for guiding, and supporting, the definition and the implementation of organizational goals and strategies.

In other words, strategic plans and decisions influence operations and activities of the organization [120], while results of the organizational activities have to reflect on organizational strategies and reshape them, if needed [62]. This is a cycle that in knowledge-creating companies or in learning organizations is characterized by meaning, management, and measurement [85]. Additionally, it has the same purpose as the
well-known PDCA\(^3\) cycle, but with an accent on organizational knowledge and learning [85]. Therefore, organizations need to understand how their goals are interconnected, how their strategies are connected to their goals, how their goals and strategies depend upon each other, and when their strategies are working and not working effectively. Those are concerns that organizations need to address, and the major concerns can be formulated as:

A. **Alignment.** Organizational alignment represents knowledge and understanding of how the different odds and ends of the organization relate to each other. Aligned organizations use that knowledge successfully to define a synchronized and consistent set of organizational goals and strategies, i.e., the organizational big picture. The creation of an organizational big picture often involves making alignment assumptions, such as assumptions about how different parts of an organization will affect joint strategies.

B. **Communication.** This is often referred to as organizational transparency or visibility. It is about making the organizational alignment visible, i.e., communicating and sharing the organizational big picture across segments of the organization.

C. **Measurement.** Once organizations have a clear big picture, or vision, of how the business should look, they need to collect evidence or data to support and guide the implementation of the organizational goals and strategies, e.g., business plans. In order to achieve this, different measurement programs across an organization need to be integrated and tailored such that they allow the organization to test the organizational alignment assumptions.

D. **Reasoning.** Collected data and measurements are used locally in projects or organizational units in which they are collected. Furthermore, organizations must reason about the organizational alignment assumptions by interpreting the collected metrics in the context of the organizational big picture, which involves different forms of analysis. For example, these forms could include the analysis of the value (cost and benefits) of selected sub-goals and strategies, their impacts and contributions to the upper-level goals and strategies, and the assessment of the risks connected to those goals, including how those risks affect the success of other goals in a grid. Such reasoning and analysis will generate new knowledge about organizational alignment.

---

\(^3\)Plan-Do-Check-Act cycle is attributed to Walter Shewhart, and it is also known as Shewhart cycle
Alignment, communication, measurement, and reasoning are elements of the organizational knowledge cycle, which will be referred as the ACMR cycle (Figure 1). Organizations constantly divide their attention between current activities, e.g. ongoing projects, which will be referred to as the operational activities, and a continuous planning of business, referred to as the strategic activities (Figure 1). The strategic activities deal with forecasts of the future and, depending on a current situation, steer the entire organization toward a new direction. For a successful steering, the knowledge about how different parts of an organization relate to each other—the alignment knowledge—is valuable and needed. The ACMR cycle starts with the alignment knowledge (A), i.e., defining and aligning organizational goals and strategies, and making that alignment knowledge visible by communicating it to the people (C). Then, the communicated knowledge, vision, and strategies are made operational, i.e., they drive operational plans and activities, and as the activities are conducted, they are measured (M), allowing the effectiveness of the activities to be evaluated and the company to reason (R) about whether these were the right goals and strategies, what the risks are in achieving the goals based upon an evaluation of sub-goal success, and which goals and strategies need to be changed/improved to achieve the high level organizational goals. All that together leads to new insights and understanding of organizational alignment (again A).

The ACMR cycle is relevant for any organization; however, this dissertation will focus on ICT and software development companies. Software companies deal with specific situations in which a true “ownership” of products, services, or business involves multiple stakeholders: end-users, customers (sponsors), creative people involved in development, business owners, etc. All of them make decisions that impact the end-product significantly [130, 150]. All of them value differently the same product or service; therefore, the entire ACMR cycle is dependent on a human perspective, the perspective of people who can perceive what is valuable for them and for their company.

1.1 Motivation and Research Opportunities

The software industry has at its disposal several approaches that accentuate the concept of value, e.g. lean development [130, 150], or agile methods [61, 76]. Still, the only prominent framework that deals with the concept of value in software development is value-based software engineering (VBSE) [36]. The VBSE framework illustrates the

---

*Information and Communication Technologies*
complexity of tasks, which are not trivial, and that are required in order to understand the value in the software development context. All those tasks require the participation of different people (stakeholders) across an entire organization and beyond the organization (customers). Only those people are capable of utilizing their knowledge and experience to make decisions that they believe are aligned with their understandings of the value.

Software engineering is a collective activity (social component), but at the same time it is dependent on individual capabilities and affinities (psychological component). Consequently, the problems that arise in software engineering are colored with socio-logical and psychological phenomena. According to Ewusi-Mensah [74] study, those phenomena have a significant impact on projects’ success rates and consequently on business success. One of the main challenges is supporting and motivating creative people to use their knowledge and experience effectively while solving problems. For example, it has been shown that the productivity of developers and programmers varies and depends upon individual characteristics, such as personality, intelligence, and expertise [102], which greatly impacts an organization’s ability to manage value creation.

One research perspective, often explored in corporative sociology and psychology, is task motivation and task performance. Edwin A. Locke, a psychologist, dedicated his lifework (since 1960s) to a theory of task motivation and goal setting—the Goal Setting Theory [135, 136]. The theory explains how goal setting impacts the motivation and performance of workers. The four mechanisms of the theory are [135]:

1. Goals serve a directive function by directing attention and effort toward goal-relevant activities, and away from goal-irrelevant activities. The effect occurs both cognitively and behaviorally.
2. Goals have an energizing function. High goals (challenging goals) lead to greater effort and attention.
3. Goals affect persistence. Work motivated by challenging goals tends to utilize available resources better.
4. Goals activate cognitive processes, and knowledge cycles, that allow employees to cope with given situations.

Furthermore, the Goal Setting Theory defines moderators. An example of a moderator is goal commitment. The goal-performance relationship is strongest when people are committed to goals [135], and self-commitment can ensure that level of a goal commitment. Self-commitment is an act of committing to self-defined goals. The concept
was first used by the team software process (TSP) [111], and was later incorporated into agile software development methodologies [61].

The Goal Setting Theory is a general theory, i.e. it does not differentiate between creative environments, like the creative industry or the software industry, and the traditional manufacturing industry. Therefore, this dissertation refers to the work of Mumford [156], who identified a special case of human intensive environments in which (creative) people work on novel problems as innovation-oriented environments. According to that separation, software development is not always innovative, if the product creation process is framed with known and existing tools and methods. Several studies reviewed by Mumford [156] suggested that externally dictated (set) goals are inhibitors of creative work, especially if the goals are evaluated against work results, which is also in agreement with the Goal Setting Theory. However, it also shows that the Goal Setting Theory has stronger implications in creative industries. Going back to the findings of the Ewusi-Mensah study and unrealistic project goals and objectives as one of the factors for failure, it can be conjectured that unrealistic goals actually cannot activate the mechanisms of the Goal Setting Theory, and that causes project failures.

Besides the monitoring function of an organization’s operations (operational activities), the measurements in the ACMR cycle have the additional purpose of collecting evidence for testing the organizational alignment assumptions, i.e. to test the hypothesis of how different units of the organization affect joint strategies and upper-level goals. In order to achieve that, different measurement programs across an organization need to be integrated and tailored in such way that it is possible to test organizational alignment assumptions. The measurement and metrics have always been perceived as important tools for learning and building knowledge, or as Lord Kelvin says:

"I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the state of Science, whatever the matter may be."

The recent development of the GQM*Strategies [17] methodology was motivated by the existing gap, or misalignment, between goal-driven, i.e. GQM³ [12, 21], measurement programs and higher-level organizational goals and strategies. The GQM*Strategies

³Goal Question Metric
represents an integrated measurement framework that helps organizations develop measurement programs capable of testing organizational alignment assumptions. The successful implementation of the GQM+Strategies methodology requires involvement of knowledgeable people from all organizational levels, and only those people know which goals and strategies are valuable for the company and how they are connected at different organizational levels.

Some approaches address parts of the ACMR cycle, e.g., value-based software engineering (VBSE) [30, 37], GQM+Strategies [16, 17], the Benefit Realization approach [187] etc. However, they fall short of implementing a complete cycle of goal alignment and sharing, measurement, and reasoning through the evaluation and improvement of organizational goals and strategies.

What follows will be a description of the research problem areas and research opportunities within those areas that motivated this dissertation work. As discussed in the previous section, the concept of value is important for software engineering because it provides a common ground to integrate different perspectives of the software development process. Furthermore, the creative people that are at the core of the software development process need to have a clear vision of organizational plans, short-term and long-term, as well as mechanisms to track their own success and its impact on higher-level organizational goals.

### 1.1.1 Value-Based Software Engineering

Boehm [36, 37, 44, 47] proposed the value-based software engineering (VBSE) framework in order to integrate all aspects of the software creation process from the perspective of value. The concept of value-based software engineering emerged in the late 1990s in the areas of product line engineering [75] and software economics [47]. The value-based approach helps focus and prioritize development efforts. Treating all requirements, use cases, and other development artifacts as equally important may have a major influence on the software project cost and schedule. As the role of software becomes increasingly important in most systems, value-neutral decisions may result in project delays, extra cost, or low quality [36].

Boehm [36] introduced seven key elements that form the foundation of value-based software engineering:

1. *The Benefits Realization Analysis* means that all initiatives need to realize whether the potential benefits of a system are identified and coordinated. Linking resources
to outcomes increases the concreteness of a software project, and helps to identify stakeholders who need to be involved in system development. The analysis also results in visible contributions, outcomes, and assumptions about the system [36].

2. **Stakeholder Value Proposition Elicitation and Reconciliation** involves identifying and documenting success-critical stakeholder value propositions. Approaches that help in agreeing upon the value propositions include expectations management, visualization and tradeoff-analysis techniques, prioritization, the use of groupware, and business case analysis [36].

3. **Business Case Analysis** involves determining the costs, benefits, and return on investment of a system during its life cycle. Unquantifiable benefits, such as stakeholder good will and uncertainties, e.g., assuming that a product will be the first of its kind in the market, make the business case analysis challenging. Analyzing uncertainties helps in identifying risks related to each development option [36].

4. **Continuous Risk and Opportunity Management** means that risk analysis and risk management should be carried out during the entire life cycle of the system. Risk management involves understanding and addressing people’s utility functions and using risk to determine how much is enough, e.g., when to stop testing, so that resources are not wasted on inessential points [36, 108].

5. **The Concurrent System and Software Engineering** stresses using iterative process models instead of waterfall style models. Concurrent process models typically have milestones that provide common commitment points [36]. For example, the eXtreme Programming paradigm is well-aligned with the VBSE approach [103].

6. **Value-Based Monitoring and Control** deals with monitoring the realization of the business value of outcomes at the project and organizational levels. It must be noted that even though a project is very successful with respect to its cost-related value, it could be a disaster from the organizational value viewpoint [36].

7. **Change as Opportunity** means that the ability to adapt to change has business value, as the rate of change is continuously increasing. Companies that can react quickly will be more successful. The philosophy is very similar to that used in agile software development [103].

The stakeholders’ viewpoints are essential in considering several of the key elements of VBSE. In addition to the reconciliation of stakeholder value propositions, stakeholders’ values are the basis of risk management, monitoring, and change management [36].

The attractiveness of concepts and ideas introduced by the value-based software engineering resulted in establishing the community of researchers and practitioners gath-
ered around these ideas and concepts. However, the extent of the software engineering field, and the verity of problems encountered in the field, provides an ample space for further research contributions to value-based software engineering.

**Research Opportunities**

Recently, Jan & Ibrar [113] conducted a systematic literature review of the VBSE field, which is a good snap-shot of the state of practice in the field. They divided the overall VBSE field into nine subfields: value-based requirements engineering, value-based architecture, value-based design, value-based development, value-based verification and validation, value-based quality management, value-based project management, value-based risk management, and value-based people management (Figure 2).

The left side of the bubble plot in Figure 2 represents contributions to each subfield. In total, they found 428 distinctive contributions. The use of metrics in the context of value-based software engineering is the least contributed topic in all VBSE subfields. Furthermore, there are no contributions regarding metrics in the subfields of value-based project management, quality management, or risk management. All those subfields are expected to utilize and rely on quantitative approaches [81, 110, 149].

**Value-based project management** Jan & Ibrar [113] found that the most common technique used for project monitoring and control is earned value management (EVM). EVM [43] techniques are used to monitor and control project performance, either for internal projects or subcontracts. One of the problems with EVM is that it can be an effective technique for managing a project’s budget and schedule, but the results of the project can still be poor from the point of view of the contributions to the organizational value, e.g., projects are completed on time and within budget, but the knowledge and experiences were not included into the organizational repository, the project results will not be reused, and the organization will lose the opportunity to convert a past project’s results into revenue. Therefore, Boehm & Huang [43] suggested the integration of critical stakeholders’ views of value with earned value analysis (EVA) through the benefits–realization approach [187] and risk/opportunity management practices. The benefits–realization approach enables identification of outcomes and assumptions related to the realization of the outcomes. A sequence of outcomes forms a results chain. The analysis of the results chain provides a basis for risk/opportunity management. Thus, a value-based version of earned value monitoring and control is performed.
by complementing EVA with risk/opportunity management and analyzing benefits realization. Benefits realizations are often not visible before the end of the project. Therefore, the value-based versions of EVM tools need to bridge the gap between actual work done in projects and benefits realizations that occur after the projects’ completion.

**Value-based risk management**  Risk management is tightly connected with project management practices. Mature organizations (e.g. organizations that are at CMMI [60] maturity level 3 or higher) treat risks with more attention and with different risk management methods and tools. The basic idea of risk management is to identify early situations or factors that may cause project problems, develop risk mitigation plans, and monitor and implement projects’ mitigation plans during project execution. Jan & Ibrar [113, p.77] found that further research is needed for developing value-based risk management methods. Traditional risk management is used by projects in order to increase the projects’ success rates by considering different types of risk. However, in the context of value-based software engineering, the value risk or the risk of failing to deliver or create value is the main concern of value-based risk management. Current project-level risk management methods do not meet the challenges of value-based risk management. Again, one of the reasons is that value materialization occurs not only during projects, but also after the project completion.

**Value-based software engineering notation**  Boehm [37] accentuated the importance of understanding the links among technical decisions, context, and value creation in different situations in order to improve decision-making. Furthermore, dynamic monitoring and control mechanisms taking into account these linkages, and different sources of value, are needed to guide decision–makers [37]. However, the use of different approaches, with different output formats, for each of the seven VBSE key elements makes the linkages less clear and less visible. Currently, there are no methodologies that can accommodate and support all seven key elements of value-based software engineering by providing a unified notation for the VBSE framework.

### 1.1.2 Goal-Driven Measurement Alignment

The issue of the alignment of organizational strategies and measurement programs in the software industry was not perceived as a research problem until the 2000s. Becker & Bostelman [23] formulated the measurement alignment as a research problem in software
engineering and accentuated a need for having common measurement frameworks in companies. They suggested a use of a combination of goal-driven measurement approaches and Balanced Scorecards (BSC) to improve the measurement alignment [23] in terms of providing better insights into corporate business and operations. The success of such organizational measurement programs depends on the abilities of organizations to align themselves. The organizational alignment represents a capability to synchronize different odds and ends of an organization to work together toward common goals. The importance of organizational and measurement alignment is accentuated by different standards, such as CMMI, COBIT, ISO family of standards, etc. Often the organizational alignment is accompanied by the concept of organizational transparency, or what is referred to as communication and visibility in the ACMR cycle (see section 1.1). Transparency is about establishing a bidirectional visibility between actions and operations at different organizational levels. In the software development context, lean software development [150, 197] considers transparency as a key to an understanding of value and waste creation in the software development processes.

From the metrics perspective, the concepts of alignment and visibility (transparency) are crucial for establishing integrated organizational measurement programs. Organizations, especially larger ones, are comprised of different units, projects, groups, and so on, with their own agendas and measurement programs. An organization-wide measurement program has to make an organization smarter and wiser about their business and operations. Usually, the integration is done in a mechanical way by providing a measurement framework, e.g. a metrics catalogue, which in turn can provide some benefits to the organization, but not a substantial insight into business and operations.

Within this dissertation’s scope are goal-driven measurement approaches, i.e. GQM, and research in the area of aligning GQM-based measurement programs with business objectives. The GQM [12, 21] is one of the most popular goal-driven measurement approaches. According to the SWEBOK⁶, the GQM is a de facto standard for goal-driven software measurements. The GQM advocates an idea of defining the measurement goals up front, followed by identifying questions to answer, and in turn the questions lead to metrics definitions [21]. This process ensures derivation of purposeful and needed metrics.

A common criticism of the GQM approach is that it is not capable of addressing corporate-level goals [55, 69, 90, 192]; and it is true, because the GQM was designed for collecting project-level data. Following that criticism two directions of research can be identified. First, some expanded the use of the GQM from the project-level to a business process-level data collection and interpretation methodology. As a result, several business process measurement frameworks and models based on the GQM were proposed [4, 87]. The second direction took an inside out view by acknowledging the existence of different groups of people inside an organization, and it uses their knowledge and experience to define an integrated measurement program. Offen & Jeffery [160] proposed the M3P measurement approach, which links GQM-based measurement programs at the project, process, and business levels of an organization. Here, the issues of alignment and visibility are critical for a successful implementation of such a measurement approach.

The GQM+ Strategies [17] is a recent upgrade of the GQM approach designed to help the software industry develop measurement programs that are aligned with business goals. The major output of applying the approach is a seamless hierarchy, called a grid. The grid integrates and aligns different measurement goals and related business and multi-level organizational goals. The first version of the method was published in 2007 as a white paper [13], followed by a series of publications illustrating the method’s usefulness on limited examples [14, 16]. Afterwards, further method development followed empirical trials with the software and ICT industries in the EU, Asia, South America, and North America [17]. The overall feedback from empirical trials accentuated the importance of the research and development of effective tools and methods for handling complex grid structures, visualizing grid structures, automating the metrics interpretation process, and providing adequate software tool support for the grid derivation activities.

The GQM+ Strategies is designed to address business information needs in a goal-driven fashion and allow business people to speak in their language—the language of business—when developing measurement programs. This characteristic of the method is needed for an approach that aims to be a framework for integrating and metricating goals from different organizational levels. In addition, only the language of business is universal enough to accomplish that extent of integration. The GQM+ Strategies elevated the GQM from a measurement paradigm for designing project-level measurement.

\(^7\)GQM+ Strategies\(^{\text{®}}\) is a registered mark of the Fraunhofer Institute for Experimental Software Engineering, Germany and the Fraunhofer USA Center for Experimental Software Engineering, Maryland.
programs (operational activities, Figure 1) to an organizational methodology for supporting the definition and implementation of goals and strategies (strategic activities, Figure 1). Research work with the GQM*Strategies is in the early phase, which is an opportunity for different contributions to this topic.

Research Opportunities

Research on the GQM*Strategies methodology is positioned within the measurement alignment research that recognizes the importance of different stakeholder groups within organizations and involves them in developing measurement programs. The advantage of this kind of approach, despite a significantly larger initial effort needed to implement such measurement programs within companies, is that it is capable of channeling organizational knowledge and establishing learning cycles. This is a significant advantage over approaches that treat (business) processes as black-boxes and measure only the top-level attributes of the processes.

Organizational reasoning about goals and strategies The GQM*Strategies is purposefully designed to address issues of organizational alignment, communication, and measurements [13, 14, 16] (ACMR cycle in Figure 1); however, organizational reasoning is partially supported through the GQM*Strategies interpretation model. Organizations can utilize different forms of reasoning, from basic reasoning about the effectiveness of strategies by analyzing the levels of goals achievement, over cost/benefit and risk analysis of the goals and strategies, up to sophisticated reasoning models for analyzing dependences between goals and strategies. Currently, the GQM*Strategies supports only the basic form of reasoning by providing feedback on goals achievement and strategies effectiveness.

Supporting methodologies The evolutionary path of the GQM approach, from its inception in the 1980s, led to the development of numerous supporting methodologies such as abstraction sheets [189], methodological packages [12], and several software tools developed to support the GQM-based measurement programs [70, 165]. It is perceivable that research in the direction of developing supportive methodologies for the GQM*Strategies approach is needed as well, especially the development of software tools for visualizing grid structures and tools for supporting the GQM*Strategies interpretation process and analysis of grid structures.
1.2 Overview of the Research Questions and Research Approach

The purpose of this introduction chapter is to present a wider problem area in which the dissertation work will be positioned. Addressing the whole problem area exceeds the scope of a single dissertation. More specifically the focus of this dissertation is on organizational reasoning and alignment (Figure 1). As shown in Figure 1, it is possible to approach the issues of organizational reasoning and alignment from two directions. First direction is from operational activities, i.e. measurement, toward organizational reasoning and alignment. The second direction is from strategic activities, i.e. from the top-management and business perspectives. The both direction are interesting from the research point of view. However, in this dissertation the first direction is taken, because the practical work in research projects was mainly concerned with measurements and how to expand the visibility of measurements toward higher organizational levels. And, for that purpose the GQM*Strategies approach was used.

The work presented in this dissertation aims at evolving and improving the GQM*Strategies approach. The research approach used in this dissertation is an adoption of the design science framework [106, 180]. An important principle of design science is that research questions have to be grounded in practical problems or needs that are identified by individuals or groups as a result of all possible forms of interactions among people, organizations, and technology.

Therefore, the first empirical investigations of the GQM*Strategies approach were undertaken (Chapter 4). The aim of the empirical investigations was to study the deployment and use of the GQM*Strategies approach in a real world setting. In particular, the goal was to look for practical problems that practitioners were addressing or trying to solve with the GQM*Strategies grid development. Those pre-studies provided valuable feedback from practitioners regarding the method’s usefulness and suggestions for further method development. That feedback helped to formulate and motivate the research questions that are posed in this dissertation (sections 4.6.1, 4.6.2, and 4.6.3). The research questions are positioned within the areas of the research opportunities discussed in sections 1.1.1 and 1.1.2.

RQ1: How can the GQM*Strategies approach be used to document value propositions across an organization? The GQM*Strategies can help in linking and connecting goals from different organizational levels. Therefore, it is an
interesting research question to explore further how the GQM* Strategies can provide a notation for enunciating an organization’s value proposition (section 4.6.1).

**Connection with research opportunities and problem areas**

This research question is directly connected with value-based software engineering notation by making clear links among various decisions across organization, context, and value propositions. Furthermore, making explicit value propositions can increase the ability of the GQM* Strategies to deal with alignment and communication parts of the ACMR cycle. This sets preconditions for a successful organizational reasoning about goals and strategies.

**RQ2: How can the GQM* Strategies be instrumented to monitor and control a value materialization process?**

Having capabilities to perform cost/benefit analysis of organizational goals and strategies was perceived as an important extension of the GQM* Strategies approach. The research question leads to the first sub-question: **RQ2.1: How can the cost/benefit analysis of the GQM* Strategies grids be performed?** Furthermore, it seems to be a very common situation in software-intensive organizations that stakeholders at one level do not have a full understanding of issues from other levels, and because of that, they are not capable of understanding the risks involved beyond their own goals and strategies. Therefore, it is important that risks are assessed when the cost/benefit analysis is performed, which is a motivation for the second sub-question: **RQ2.2: How can risks regarding the value materialization be identified?** (section 4.6.2)

**Connection with research opportunities and problem areas**

This research question is directly connected with value-based project management and value-based risk management. The earned-value management approach is extensively used for value-based project management (section 1.1.1). However, it fails short in addressing properly the benefits realization from projects. In the ACMR cycle, the abilities to monitor and control value realization, and associated risks, involve the organizational measurement and reasoning activities.

**RQ3: How can the GQM* Strategies be used to analyze dependences between different value propositions across an organization?**

Different strategies can have different effects on a same value proposition, i.e. they can lead to different ways of value realization. Some value propositions resulting from one strategy can even have a
negative effect on other value propositions, which is usually manifested as contradicting goals. Therefore, it is important to provide a mechanism for analyzing and quantifying goals-strategies dependences: **RQ3.1**: *How can the goals-strategies dependencies in the GQM+Strategies grids be quantified?* Furthermore, such a mechanism could be used to further analyze the impacts of risky goals on other goals and strategies in the GQM+Strategies grid. **RQ3.2**: *How can the impact of risky value propositions on other value propositions be assessed?* (section 4.6.3)

**Connection with research opportunities and problem areas** The focus of this research question is on *organizational reasoning about goals and strategies*. The GQM+Strategies approach provides some basic reasoning mechanisms, but it does not support the analysis of the value (cost and benefits) of selected sub-goals and strategies, their impacts and contributions to the upper-level goals and strategies, and the assessment of the risks connected to those goals, including how those risks affect the success of other goals in a grid (section 1.1.2). Generally, this research question focuses on organizational reasoning and abilities to perform different forms of analysis over grids, i.e. the reasoning part of the ACMR cycle.

All research questions are practically oriented, i.e. asking *how* to utilize the approach for a specific problem. Such questions call for a further development or evolution of the GQM+Strategies approach. According to the design science research framework [106], a solution development represents a form of a search process. That process is characterized by design decisions that have impact on the solution. In order to make those decisions explicit, the *design hypotheses* are formulated. For the purpose of making a clear distinction, those hypotheses are not tested or evaluated directly, like in some other research approaches. However, once a solution is developed, the design hypotheses can be evaluated to determine if they were right or if some other hypotheses would lead to a better solution.

The research on solution development utilized analytical and empirical paradigms in different phases of the solution development. The analytical paradigms were used for solution development and evaluation, while empirical paradigms were used for evaluating certain aspects of the solution.
Overall VBSE
23

8

22

5

5
1.17%

1.87%

21

Value-based
Requirements
Engineering

1.17%

5.14%

5.37%

4.91%

13

53

9

6

2.96%

2.05%

1.37%

5
1.14%

12.07%
4

3

7

0.93%

0.70%

1.17%
11

2
2.57%

4.21%

30

2
0.47%

6

9

1.17%

2

1

0.23%

0.47%

8
0.47%

Value-based
Design

1.17%

1
0.47%

2

5

5
0.23%

2

Value-based
Architecture

2.10%

1

0.47%

2.10%

9
1.40%

1
0.23%

5

18

1.64%

10
1.87%

0.23%

2
2.34%

0.47%

Value-based
Development

Value-based
Verification &
Validation

8

20

5

3

1.82%

4.56%
3

2
0.68%

10
1.37%

4.10%

3

14

10
0.46%

0.68%
4

3
3.19%

0.68%

27

4

2.28%

20
4.67%

9

1

1
0.23%

2.10%

4.67%

15
0.23%

3.50%

Value-based
Quality
Management

0.91%

4
0.91%

0.91%
6.15%

7.01%
20

0.23%

0.68%

2.28%

2

18

0.68%

3
0.46%

6

1

1.14%

31

16

7

15
1.59%

3.64%

3.42%

2
0.46%

7.06%
13

8
1.87%

3.04%

7

11

7
1.64%

12

10

6
0.70%

2
0.47%

19
1.40%

3
2.34%

2
0.23%

6
0.70%

2.57%

1.64%

1
2.80%

3

15
1.40%

4

0.47%

Recommenda
Method/
Model
Tool
tion/practice 67(15.65%) technique 20(4.67%)
132(30.84%)
69(16.12%)

4.44%

3.50%

3
0.93%

Framework
38(8.88%)

0.70%

Process/
Metric
Approach 3(0.70%)
99(23.13%)

Value-based
Project
Management

Value-based
Risk
Management

Value-based
People
Management

15

24

8
1.37%

2
1.82%

0.46%

5.47%

14

28

1

3
0.23%

3.19%

1
0.68%

0.23%

6.38%
11

3
2.51%

Evaluation
Research
171(38.95%)

Value-based Contribution
428 (100%)

Fig 2. Bubble plot of overall VBSE. Source: [113, p.38].

5
0.68%

11
1.14%

2.51%

Opinion
Solution
Experience Validation Philosphical
Report
Research
Report
Research
Research
2 (0.46%) 111 (25.28%) 52(11.85%) 70(15.95%) 33(7.52%)
Research Type
439 (100%)

F igure 6: Bubble plot of Overall VBSE

36

6

3.42%


1.3 Terminology

In this section, the key terms used throughout the dissertation are explained with the intention to set up a common understanding.

**Organizational Goals**, or simply **goals**, are statements of objectives. They can be expressed in different ways, from verbal statements to formally written definitions. They help to focus attention on relevant issues and impediments regarding the achievement of goals. For example, even an informal discussion or brainstorming about goals is accompanied by considerations of constraints, feasibility, and related issues. Therefore, the references to goals, besides the objectives that need to be met, also refer to the relevant issues that need to be addressed. The prefix **organizational** is an umbrella term used with the intention to include goals from different levels (e.g., project goals, software goals, quality goals), and not only the top-level business goals.

**Strategies** represent plans to solve identified issues and to achieve objectives that are defined with organizational goals. There is a vast literature on strategy formulation and management. However, there is now clear consensus about what a strategy is [28], but rather a common understanding of a strategy as a guiding mechanism for setting goals and conducting activities in an organization. The diversity of writings on this topic is well summarized by Mintzberg *et al.* [153] who identified ten different strategy schools: (1) The Design Schools, (2) The Planning School, (3) The Positioning School, (4) The Entrepreneurial School, (5) The Cognitive School, (6) The Learning School, (7) The Power School, (8) The Cultural School, (9) The Environmental School, and (10) The Configuration School. The first three schools are **prescriptive**—concerned with how strategies should be formed—and they are among the first schools. The next six schools are more **descriptive**—concerned with the process of strategy formation. The last, tenth school, is trying to combine various elements of the previous eight schools. To this list of strategy schools, recently after the economic crash in 2008, the Pragmatic School resurrected with their ideas of pragmatism in strategy formulation [159]. The pragmatic strategy [159] is very close to the Cognitive and Cultural Schools. All these schools have different views on strategy formulation and definition process. For example, some schools advocate simple and informal process of strategy formulation (e.g. Design and Learning Schools), while some schools are very systematic and formal (e.g. Planning and Positioning Schools). Throughout the dissertation the concept of strategy in a
generic form is used. Thus, a strategy can appear at any corporate level (e.g., strategic, tactical, operational) wherever organizational goals are defined. Also, strategies can be defined as statements that later can evolve into detailed plans with an activity-like structure.

**Stakeholders** as a concept originates from Freeman’s work on Stakeholder Theory [82]. Since then, it is used to refer to people, either individuals or groups (e.g., organizational units, social structures) whose understanding and perception of what value is, is relevant and important [154]. While in law, a stakeholder represents a party or a legal entity that holds some rights or property. Traditionally, the focus of stakeholder value analysis used to be on customer and business value; in other words, the perspectives of customers and business owners were dominant. Nowadays, in knowledge companies respecting creative people at various organizational levels has became a necessity [158, 159], so their perception of value is relevant as well. And, this definition of stakeholders maps well onto the adoption of the stakeholder concept by the value-based software engineering community [38, 40, 44, 97].

**Goal Owners** are representatives of a stakeholder population who are involved in defining organizational goals. This term is introduced with a purpose of linking stakeholders and GQM+ Strategies approach. In an idealized situation, the goal owners would be identified by their involvement in the process of defining and refining organizational goals and strategies. However, in practice the goals ownership is often a given responsibility by superiors, who give explicit tasks to their subordinates to participate in the goal and strategies definition and refinement process. As a result of identifying goal owners, the goal’s commitments [135] are clear and visible.

**Value** as a term is overloaded and used in different contexts throughout the literature. Value as a concept exists as the result of humans’ psychological and social affinities. A person who values certain goods more highly than other goods is consequently willing to pay more for the former. Therefore, value is always subjective, and cannot be separated from the stakeholders’ viewpoint. Such definition of the value is widely accepted by economics through the Utility Theory [79, 185], which suggests that people have a preference toward options that can gain more benefits. However, according

[^8]: http://encyclopedia.thefreedictionary.com/Stakeholder+(law)
to Biffl, Aurum, Boehm, Erdogmus & Grünbacher [29], the expected benefits can be tangible or intangible, economic or social, monetary or utilitarian, or even aesthetic or ethical. In other words, the value is a multidimensional concept that admits multiple characterizations chosen or set by the value beholder. And, analyzing all these qualities of the value (i.e. that someone deems valuable) falls into the domain or real options theory, which is an extension of financial option theory to options on real (non-financial) assets [186]. In order to accentuate the subjective component of the value concept, the prefix perceived is used, referring to stakeholder’s opinion and belief about the value. In the context of this dissertation perceived value is also used to accentuates an important dimension of the overall concept of value—the temporal context. Value as a phenomenon is bound by time; that is to say, stakeholders’ understanding of value changes through time. The most important aspect of time in relationship to value is the future, which impacts what is referred to as perceived value. Perceived value reflects stakeholders’ beliefs about what will be valuable at some point in the future, e.g., when defining organizational goals that have to be achieved in a certain timeframe. This is a type of value that will be dealt with throughout this dissertation, and very often the term perceived value will be used without its prefix perceived.

**Value Proposition** is a set of objectives or requirements that stakeholders would like to meet or satisfy [97, 100]. Also, in the context of this dissertation the value proposition represents stakeholders’ belief why certain strategies will be successful and beneficial for an organization. Making value propositions explicit involves a process of clarifying and negotiating the propositions among stakeholders [100]. This is the definition of value proposition used by value-based software engineering community.

**Alignment** is defined by the dictionary\(^9\) as the act of aligning or state of being aligned; especially: the proper positioning or state of adjustment of parts (as of a mechanical or electronic device) in relation to each other. Organizational alignment is constantly among the top-ten concerns of corporate executives in the last two decades [138]. The importance of organizational alignment is well addressed by management and business literature [49, 119, 120, 138, 182]. However, as noticed by Kaplan & Norton [120], aligning organizational units to create value at the enterprise level gets less attention, in theory and practice, than creating value at business unit level. For a

\(^9\)http://www.merriam-webster.com/dictionary/alignment
corporation to add value to its collection of business units and services it needs to align those units and create a synergy, i.e. a corporate added value [120]. One commonly researched topic is strategic business-IT alignment [105, 137, 184] that focuses on the support of information technologies for a successful implementation of business objectives and plans. Some authors argue that alignment permeates an organization at all levels [49, 120] and it plays an important role in synchronizing organizational activities. For example, Boswell & Boudreau [49] explain this through the concept of the line of sight. The employees with the line of sight to organizational objectives contribute better to organizational strategies. Furthermore, Boswell & Boudreau identified four types of employee understanding of strategic organizational objectives, i.e. characterization of the organizational line of sight: (1) accurate and deep, (2) accurate and shallow, (3) inaccurate and deep, and (4) inaccurate and shallow. People with accurate and deep understanding do not only understand the strategic objectives, but also see how they help the organization to achieve those objectives [49]. Therefore, when the term alignment is used in dissertation, it refers to a holistic view of how different pieces form a big picture. Each piece (e.g., project goals, unit goals and plans, corporate goals and strategies) carries relevance and significance, but at the same time having an understanding of how pieces relate to each other is important for the successful steering of an organization toward business success.

**Reasoning** is defined by the dictionary\(^\text{10}\) as *the drawing of inferences or conclusions through the use of reason [argument]*. People can use different forms of reasoning or sense making, from a free form argumentation up to formal logical systems. In this dissertation, when the term organizational reasoning is used, it refers to all different methods and techniques that organizations can use to reason about their goals and strategies.

**ICT Industry** represents, in a very generic form, businesses that are providing solutions based on information and communication technologies (ICT\(^\text{11}\)). That includes organizations that have for their main product ICT-based solutions (products and/or services), but it also includes departments within larger corporations that deal with ICT problems, i.e. IT departments. The purpose of referring to the ICT industry, in this generic form, is not to scope down or limit the impact of this dissertation, but rather it

\(^{10}\)http://www.merriam-webster.com/dictionary/reasoning
\(^{11}\)http://encyclopedia.thefreedictionary.com/ICT
came as a joint characteristic of the organizations involved in empirical studies that were conducted during the dissertation work. Otherwise, the targeted population that can have benefits from this work can be characterized as software-intensive or even knowledge-intensive industries [158].

### 1.4 Notes on the Dissertation Structure

The dissertation structure reflects the research framework that was used. Chapters are logically grouped in four parts with an introduction (Figure 3).

<table>
<thead>
<tr>
<th>PART I: (Ch 2 and 3)</th>
<th>PART II: (Ch 4 and 5)</th>
<th>PART III: (Ch 6, 7, and 8)</th>
<th>PART IV: (Ch 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Related Work</td>
<td>- Pre-study</td>
<td>- Organizational Value Alignment and Measurement</td>
<td>- Results and Discussion</td>
</tr>
<tr>
<td>- Background Technology</td>
<td>- Research Questions</td>
<td>- Organizational Earned Value</td>
<td>- Solution Validity</td>
</tr>
<tr>
<td>- Research Approach</td>
<td>- Research Problem and Design Hypotheses</td>
<td>- Organizational Causal Model</td>
<td>- Future Work</td>
</tr>
</tbody>
</table>

In the introduction chapter the research problem areas (sections 1.1.1 and 1.1.2) are introduced and motivated. In Part I, related work is discussed (Chapter 2) and the research approach is presented (Chapter 3). As a part of the related work, the underlying methodology (i.e. GQM+Strategies) that is used for developing the solution is presented.

Part II presents empirical studies with the GQM+Strategies methodology (Chapter 4) that were conducted with the purpose of getting first-hand experience with the methodology to better understand the challenges with the method’s use. The feedback from empirical studies was used to formulate research questions and later to formulate the research problem for which the activities will be undertaken to develop a solution (Chapter 5).

Part III represents the solution development process. Each of its three chapters addresses a specific subproblem. Chapter 6 deals with organizational value alignment.
and measurement, Chapter 7 defines organizational earned value metrics, and Chapter 8 introduces causality theory and builds formal goal-strategies-goals models.

In Part IV, the dissertation is summarized by revisiting the research questions. The results of the dissertation and contributions are presented, followed by a discussion of validity issues. Finally, the future research directions are presented (Chapter 9).

Throughout the dissertation parts of one hypothetical example are used wherever it is convenient. The whole example is given in Appendix 2. Other appendices include interview scripts (Appendix 1 and Appendix 4), one survey (Appendix 3), and a summary of field notes (Appendix 5).
I Related Work and Research Approach
2 Related Work

The learning and knowledge that we have, is, at the most, but little compared with that of which we are ignorant.

by Plato
(427BC–347BC)

As stated earlier in section 1.1, the work presented in this dissertation is positioned in the areas of goal-driven organizational measurement alignment and value-based software engineering. The related work that is directly used for developing a solution is reviewed in sections 2.3 (GQM\(^*\)Strategies approach) and 2.4 (Value-based software engineering).

In section 2.2 an overview of the goal-driven measurement approaches is presented. Some issues identified during various applications of the goal-driven approaches motivate the development and further evolution of the GQM\(^*\)Strategies approach. The related work reviewed in sections 2.5 (Design science) and 2.6 (Empirical software engineering methods) is used for designing the research approach of the dissertation.

In what follows, an overview of the literature sources and review methodology is elaborated.

2.1 Literature Sources and Review Methodology

Literature review procedures and protocols are mechanisms for searching bodies of knowledge, i.e. literature sources. Nowadays, researches have different scientific databases at their disposal (e.g., Springer Link, IEEE Xplorer, ACM, ISI Web) for searching the literature. At the same time, the selection and use of a database is perceived as a validity issue for literature review findings. Recently, several studies were conducted to compare the quality of the search query results between Google Scholar and other electronic library databases [52, 195]. Walters [195] showed that Google Scholar performs better than the compared academic subscription databases. In other words, Google Scholar is the world’s best index library and is sufficient in most cases because it indexes all major subscription databases, and not using it is a serious threat to the validity of the findings. Therefore, Google Scholar was used as the primary indexing database. The tool used for searching references was Harzing’s PoP,\(^{12}\) which

\(^{12}\)Harzing’s Publish or Perish (PoP): http://www.harzing.com/pop.htm
uses the Google Scholars index database [195]. All references selected and used for this dissertation work comprise the dissertation’s reference database. In fact, an MS Access application was developed for storing and commenting on selected references. Such an application allows the generation of different reports, for example, Figure 4 shows the library sources and types of references in the database. Individually, the largest source of references was IEEE Xplorer (with 23%), which is expected because all major conferences in software engineering are indexed or sponsored by IEEE. Conferences that were co-sponsored by ACM and IEEE, were labeled as IEEE references in Figure 4.

Fig 4. The literature sources and types of references used in the dissertation.

The types of reference sources used by researchers can be categorized as primary, secondary, and tertiary [188]. The primary sources are data collected by researchers through empirical studies, while secondary sources represent published results (e.g. conference and journal papers). The tertiary sources are books and encyclopedias written for a wider audience. Publications that are peer reviewed, like conference papers and journal articles, can take some time until they are published and made publicly available. Therefore, in fields like software engineering where technology changes with increasing rates, conference publications are important because they have much shorter reviewing and publishing times than journals. Besides the primary sources, researchers should use more conference and journal publications than published books or encyclopedias (tertiary sources) as literature sources [188]. Figure 4 shows that journal and conference papers are dominant, making up 64% of all references used in this dissertation.

The majority of the references were identified using one of the two literature review protocols, as explained below. Besides the references that were selected through the
literature review, references that were recommended or suggested for this study were also included (e.g. suggested by relevant peers as feedback in conferences). Such references are small in number, but they tend to be of the highest quality with respect to their relevance to the dissertation work.

2.1.1 Reference-Based Literature Search Protocol

The reference-based search strategy [54] is particularly useful for evaluating the impacts that some contributions made over time, or for reviewing a historical perspective of some methods or approaches of interest. The main idea of a reference-based search strategy [54] is to utilize search engine tools for accessing citation-related information (i.e., citing papers, number of citations). A particularly important step in the reference-based search strategy is the identification of cardinal papers (CP). Cardinal papers [54] are sources that have high impact in a given research field. Such papers are referred to as well-known because they are cited often in literature.

Fig 5. Literature review protocol based on reference-based search strategy.

The reference-based literature review protocol that was adopted has four phases (figure 5):

- Phase 1. In the starting phase a list of cardinal papers is specified. This phase differs from a starting step described by Budgen et al. [54]. Experts in the field are consulted to point out cardinal papers or well-known papers rather than picking up a take-off paper [54] that references the cardinal papers.
– **Phase 2.** This phase relies on search engine tools to select *all papers*\(^{13}\) that cite cardinal papers. The number of these references correlates with the impact of the cardinal papers; for the papers with higher impact the number of citing papers is large.

– **Phase 3.** In the third phase, a filtering of the references is done by relevance criteria. In the case of this dissertation, the relevance criteria is citation of cardinal papers as a "background and related work."

– **Phase 4.** Finally, papers are filtered by reviewing abstracts, and if abstracts are inconclusive then introductions are checked. This ensures that only those papers that are addressing, validating, or evolving results of the cardinal papers are selected.

References selected in this way have a strong relation with contributions presented in cardinal papers, and as such they can be used to analyze different issues. For example, they can help determine what issues were identified with a proposed methodology (cardinal paper) and how later contributions addressed those issues.

### 2.1.2 Systematic Literature Review Protocol

For reviewing the related work, a systematic literature review protocol was also used. The systematic literature review protocol is not so formal and extensive as suggested by Kitchenham [125], but rather it is a continuous process that is run whenever new relevant keywords are identified.

The overview of the protocol is given in Figure 6, and it has the following phases.

– **Phase 1.** First, a query over the index database is defined. This is a search by keywords. The query result set contains thousands of references that are passed into the second phase. From that query result, a set of references can be selected by ranking criteria or by citation impact metrics. When a query has a complex form (i.e. multiple key words with use of logical operators), then ranking is a good indicator of the relevance of a reference. However, when a query is very simple, e.g. the name of a specific method or concept, then impact metric is a much better selection criteria, because it indicates that the language and terminology used in such references are shared among a larger audience.

– **Phase 2.** In the second phase, a filtering of the references is done by relevance criteria. The relevance criteria can be checked by reviewing the abstract and analyzing if the reference really addresses the issues of interest.

\(^{13}\)All papers that are indexed by search engine databases.
Fig 6. Literature review protocol for the related work.

- **Phase 3.** Phase three represents references that passed the previous steps. Reviewing content of those references might lead to a new set of keywords or a list of relevant authors that need to be considered. Therefore, there is a feedback loop back to phase one.
- **Phase 4.** Finally, the remaining papers represent references that are considered for the literature review. The references are filtered by content. General quality attributes are checked, like how well a paper is written.

This literature review protocol relies on the utilization of search engines, primarily Google Scholar. References selected in this way have strong correlation with search queries. The queries can be modified and changed whenever new keywords are discovered; therefore, this protocol can be seen as a continuous process of searching the literature.

The reference-based search strategy was used for reviewing the evolution of goal-driven measurement approaches. The remaining related work was reviewed using the systematic literature review protocol.

### 2.2 Goal-Driven Alignment and Measurement

The baseline methodology used in this dissertation (i.e. GQM+Strategies) is an evolution of the goal-driven measurement toward an approach capable of dealing with organizational alignment issues in a goal-driven way. Therefore, it is interesting to
introduce the following two areas of research: goal-driven alignment approaches and goal-driven measurement approaches.

Later in Section 2.2.3 a link and connection between these two areas is discussed, which will serve as a motivation for the introduction of the GQM+Strategies approach.

### 2.2.1 Review of Goal-Driven Alignment Approaches

Business and organizational alignment is one of the top concerns of IT executives in the past decades [138]. There are many contributions that investigate: different alignment frameworks, e.g. [32, 66], the role of IT in strategic alignment, e.g. [105], models for assessing alignment maturity, e.g. [137], or use of strategic alignment as a model for organizational transformation [104]. Within the scope of this dissertation are goal-driven approaches that are or can be used to deal with strategic alignment of an organization.

According to Babar et al. [5], the use of goal-oriented approaches to model strategic alignment is a novel approach, at least in IS/SE research. Interestingly, all those approaches are mainly coming from the requirements engineering domain [5, 63, 145, 202]. Babar et al. [5] conducted an evaluation of the goal-driven approaches for modeling strategic alignment concept. In order to compare different approaches, they derived 12 characteristics or requirements of an effective approach for modeling the concept of alignment (Table 1). The reference model for deriving requirements was Strategy Maps [119], a tool used by popular Balanced Scorecard approach [116, 118].
Table 1. A list of requirements for the concept of strategic alignment. Taken from [5].

<table>
<thead>
<tr>
<th>Req. ID</th>
<th>Requirement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>Stakeholders (such as individuals, departments, systems) related to the objectives in four perspectives.</td>
<td>Does the goal-oriented approach support modeling stakeholders relevant to the concept of strategic alignment?</td>
</tr>
<tr>
<td>R2</td>
<td>Intentions of the stakeholders</td>
<td>Does the goal-oriented approach support identifying strategic intent/s associated to the stakeholders?</td>
</tr>
<tr>
<td>R3</td>
<td>Customer value propositions (CVPs)</td>
<td>Does the goal-oriented approach support modeling CVPs in terms of quantitative and qualitative values?</td>
</tr>
<tr>
<td>R4</td>
<td>Dependency relationships among the stakeholders</td>
<td>Does the goal-oriented approach support modeling dependency relationships among the stakeholders for strategic targets?</td>
</tr>
<tr>
<td>R5</td>
<td>Goals (quantitative)</td>
<td>Does the goal-oriented approach support modeling quantitative objectives in strategic alignment?</td>
</tr>
<tr>
<td>R6</td>
<td>Goals (qualitative)</td>
<td>Does the goal-oriented approach support modeling qualitative objectives in strategic alignment?</td>
</tr>
<tr>
<td>R7</td>
<td>Assets (Tangible and intangible)</td>
<td>Does goal-oriented approach support modeling assets in terms of tangible–S/W and/or H/W and intangible–information entities of strategic alignment?</td>
</tr>
<tr>
<td>R8</td>
<td>Processes used to deliver financial and customer value targets</td>
<td>Does the goal-oriented approach support modeling processes in terms of activities and tasks?</td>
</tr>
<tr>
<td>R9</td>
<td>Refinement of goals</td>
<td>Does the goal-oriented approach support refinement of strategic-level goals to detailed-level goals?</td>
</tr>
<tr>
<td>R10</td>
<td>Rationale of strategic intent/s</td>
<td>Does the goal-oriented approach support facilitating rationale behind strategic intent/s of the stakeholders?</td>
</tr>
<tr>
<td>R11</td>
<td>Reasoning about strategic intent/s in terms of alternatives, contribution links and decomposition links.</td>
<td>Does the goal-oriented approach support modeling different types of relationships among the goals and stakeholders with meaningful constructs?</td>
</tr>
<tr>
<td>R12</td>
<td>Alignment perspectives</td>
<td>Does the goal-oriented approach support modeling different perspectives of an IS?</td>
</tr>
</tbody>
</table>
Babar et al. [5] examined the following five approaches:

**i* Approach.** The i* approach emerged in the area of requirements engineering [5, 202], and it is widely used to support goal-driven requirements engineering practices [145]. The approach defines two groups of concepts: (SD) Strategic Dependencies, and (SR) Strategic Rational. The SD includes concepts of actors (stakeholders of the system) and dependencies among actors. While concepts of goals, soft goals, tasks, and resources are considered as the SR. The i* notation uses decomposition links to connect and link all elements (goals, tasks, actors, etc.) that are needed for achieving a specific task.

**NFR.** Non-Functional Requirements (NFR) approach is designed to support modeling of the non-functional requirements [5]. NFR approach defines a quality goal graph, in which high-level quality goals are decomposed into more specific goals until a satisfying solution is reached. In order to support the decomposition process, NFR defines concepts of *alternatives* and *means-ends* that help in analyzing different ways of achieving goals.

**KAOS.** Knowledge Acquisition in Automated Specification (KAOS) is a formal goal modeling approach that uses graph theory to represent relationships between the goals [5, 145]. KAOS is used in later phases of requirements engineering, when a sufficient set of stable requirements is identified. KAOS supports formal reasoning and allows to automatically generate requirements specification that satisfy high-level goals. In addition, it allows to formally verify that the goals have been achieved with the use of proof theory of temporal logic.

**GBRAM.** Goal Based Requirements Analysis Method (GBRAM) represents a set of practical guidelines for the identification and analysis of the organizational goals [5]. It is also used in the later stage of requirements engineering. The method uses heuristic rule to identify goals from text analysis of various documents, like specifications. GBRAM is a prescriptive language and it does not offer any graphical notation to represent requirements. It mainly offers identifying goals and agents in a textual form which are generally organized around system goals [5].
Map. Map presents a simple goal modeling structure during requirements engineering using only two concepts of intention and strategy []. Map claimed to be an approach that uses simple notations to address complex systems which helps analysts to understand the models easily. A map provides a representation of a multi-faceted purpose on a non-deterministic ordering of intentions and strategies. A Map represents a graph with intentions as nodes and strategies as edges. An intention is an optative statement that expresses what is wanted. While, a strategy is an approach, a manner, a means to achieve an intention.

Table 2. Effectiveness of goal-oriented approaches for modeling strategic alignment.

<table>
<thead>
<tr>
<th>Requirement</th>
<th>i*</th>
<th>KAOS</th>
<th>NFR</th>
<th>GBRAM</th>
<th>Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1: stakeholders</td>
<td>√</td>
<td>√</td>
<td>x</td>
<td>√</td>
<td>x</td>
</tr>
<tr>
<td>R2: Intention of stakeholders</td>
<td>√</td>
<td>√</td>
<td>x</td>
<td>√</td>
<td>x</td>
</tr>
<tr>
<td>R3: Customer value proposition</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>R4: Dependency relationships among stakeholders</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>R5: Goals (quantitative)</td>
<td>√</td>
<td>√</td>
<td>x</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>R6: Goals (qualitative)</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>R7: Assets (i.e. tangible, intangible)</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>R8: Processes (i.e. tasks, activities)</td>
<td>√</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>√</td>
</tr>
<tr>
<td>R9: Refinement of goals</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>R10: Rationale</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>R11: Reasoning (i.e. alternatives, contribution link and decomposition links)</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>x</td>
</tr>
<tr>
<td>R12: Alignment perspectives</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

The summary of the evaluation findings is given in Table 2. All compared approaches support the concept of goal modeling and goal refinement process. However, one finding was that none of the approaches support modeling of different alignment perspectives [5], where the alignment perspective criterion is derived from the strategy map’s ability to analyze causal relationships among different Balanced Scorecard perspectives. And, this particular criterion is the only one that includes any form of identifying causal relationships. Meaning that none of the upper mentioned goal-driven approaches explicitly uses causal analysis to evaluate the alignment of the goal hierarchy.

2.2.2 Review of Goal-Driven Measurement Approaches

Contributions in the area of software measurement and metrics date back as early as the 1950s. The earliest reference retrieved in the literature review was a conference
paper on an early version of product metrics by Rubey & Hartwick [174] published in 1968; interestingly, the authors did not use any references in their paper. The book *Software Metrics* by Gilb [88] was considered as a representative reference for the early measurement efforts in software engineering. The focus of the early measurement efforts was mainly product metrics, with the exception of Boehm [35], who was among the very first authors to address process-related metrics for the purpose of developing a software cost prediction model. Additionally, Gilb [88] discussed both product metrics and productivity-related metrics.

This is not to imply that there were no other contributions beside Gilb’s work. On the contrary, software metrics were very actively studied in connection with managing the quality of software products [146–148]. The main carriers of those initiatives were governmental institutions in the United States. The majority of those works’ results were published as internal technical reports. For example, the technical report from 1977 by McCall *et al.* [147] was pointed out by Card [57] as the first publication on the goal-driven measurement approach. However, after careful analysis of McCall *et al.*’s report, it seemed very unlikely that an outside reader who was not familiar with goal-driven measurement could recognize the goal-driven principles in quality factors and related criteria, as was suggested by Card. Certainly, they were not articulated as clearly as in the first conference paper that used goal-driven data collection methodology published in 1981 by 22 [1]. This methodology was named the GQM (Goal/Question/Metric) approach [21].

With further development of the software process models and quality standards [59, 109, 149] the issue of measurement alignment became increasingly important, and it motivated further development of goal-driven measurement approaches.

The Goal/Question/Metric (GQM) [12, 18, 21] approach originated from the work of Basili & Weiss done at the Software Engineering Laboratory (SEL) at the beginning of the 1980s [21]. Even though the method was originally developed to build baselines in the SEL and for other project management purposes, it was quickly recognized by the software industry as a viable solution for establishing measurement programs.

The GQM paradigm is a mechanism for defining and evaluating a set of operational goals using measurement. It represents a systematic approach for tailoring and integrating goals with models of the software processes, products, and quality perspectives of interest, based on the specific needs of the project and the organization [12, 18]. The main concepts of the GQM paradigm are goals, questions, and metrics (Figure 7). Goals document objectives of measurements, and they are documented with the goal template
The template specifies the following dimensions of the GQM goal: object of study, purpose, quality focus, viewpoint, and context. Further on, a set of questions operationally define a measurement goal. Answers to the set of questions should clearly indicate if the goal is achieved or not. Finally, a set of measures is associated with every question to answer it in a quantitative way.

Fig 7. Illustration of the GQM graph and abstraction sheet.

The GQM paradigm advocates top-down metrics derivation, with the purpose of ensuring a meaningful metrics definition. The interpretation process goes bottom-up (Figure 7). One of the benefits of the goal-driven measurement approach is the collection of needed data with the possibility to reuse metrics and questions for different GQM goals.

The GQM approach package includes a GQM abstraction sheet [189], a tool for the derivation of questions and metrics from GQM goals. The tool is used during interviewing sessions or group discussions. One of the main advantages of the abstraction sheet tool is its ability to expose implicit models and to formulate them into questions and metrics [51].

The structure of the GQM abstraction sheet is depicted by four quadrants and a header (Figure 7). The header represents the GQM goal with five dimensions: object, purpose, quality focus, viewpoint, and environment, while the quadrants [189] represent: (I) variation factors, (II) quality focus, (III) baseline hypothesis, and (IV) impact of variation factors.

Interviewees are asked to propose relevant attributes to describe the goal’s quality focus (second quadrant). Further on, the interviewees are expected to hypothesize values of the quality focus attributes (third quadrant). After that, a discussion of variation factors and their impact on the baseline hypothesis takes place (first and fourth quadrants). If
needed changes in quality focus and the baseline hypothesis can be made, as the entire process is iterative.

Because the interpretation process is seen as structured discussions and dialogs between the GQM measurement team and others (developers, managers, testers, analysts, etc.), it is important to expose implicit models at the beginning when specifying the data to be collected. According to van Latum et al. [189], that dialog or feedback is the key factor for the successful implementation of the measurement program and later process improvement.

Research in the area of goal-driven measurement was very vivid in the 1990s. Several other goal-driven approaches were suggested, like the ami\textsuperscript{14} method [129], the GQ(I)M\textsuperscript{15} approach [164], GQMCAI that adds on existing three-tier structure layers for collection, analysis, and implementation [93], V-GQM (Validating GQM) that adds a process for defining the measurement life cycle as inspired by the V-model [163], or GAM [65] that extends the GQM with argument structures for an improved goal refinement.

The GQM approach proved to be very effective in developing measurement programs that are aligned with the objectives of the measurement, as it was used in numerous industry applications of the approach [31, 67, 92, 121, 128, 189, 190]. Also, the GQM is used as a research instrument for collecting data during empirical studies in software engineering [128, 143, 157].

2.2.3 The Gap between Measurement and Organizational Goals

The importance of the organizational ACMR cycle was discussed in chapter 1. If that cycle of alignment, communication, measurement, reasoning, and again alignment is broken, it creates a gap between organizational strategies and operation-level activities, and the measurement programs of such organizations are usually misaligned with business objectives (Figure 8).

\textsuperscript{14}ami – Assess, Analyse, Metricate, Improve

\textsuperscript{15}Goal Question (Indicator) Metric
The alignment of the measurement programs with measurement goals is important, but it is not sufficient if those measurement goals are not aligned with organizational or business goals. This was addressed by literature on multiple occasions under the issue of defining “right” measurement goals. The right measurement goals should coexist with the identified information needs of a project or an organization [9, 99, 124, 164]. Also, the goals should not be too general or too localized [9, 93, 124, 139]. If goals are too general, they will require larger sets of complex questions that will result in less reliable interpretation of such goals. An opposite situation is when goals are too localized, which could lead to a trivial interpretation. This can occur if measurement goals are misaligned with organizational and business goals.

The idea of having several types of goals has been present since the early papers on the GQM. In particular, it was suggested to have two types or categories of goals; for example, Kuntzmann-Combelles [129] introduced primary goals in addition to measurable goals in their version of the goal-driven measurement approach. The
majority of the references suggest that other goal types have to be associated with stakeholders and business objectives [23, 55, 69, 90, 192], which can be seen as ways to bridge the gap between strategic objectives and measurement systems as pointed out by Becker & Bostelman [23]. All those approaches have in common the idea of linking operational-level measurement schemes with business goals and objectives through some kind of business goal "refinement" process.

Offen & Jeffery proposed a slightly different solution for the misalignment problem. They proposed a mate-model $\textit{M3P}$ (Model, Measure, Manage Paradigm) [160] that extends the QIP/GQM paradigm by making explicit and linking organizational and business issues and concerns. They started from the fact that organizations have a layered structure, i.e. levels (strategic, tactical, and operational). For each level, a measurement scheme is developed to address information needs of that level. The linkages between levels are made thorough integrating measurement schemes [160]; however, the organizational and business goals at different levels are not explicitly linked. The $\textit{M3P}$ model is of significance because it made clear that it is not sufficient to establish measurement programs exclusively for the operational-level issues. The operational-level issues have impacts on the higher organizational concerns, but establishing measurement programs only for the operational-level issues cannot provide sufficient information for managing organizational higher-level issues.

The industry-oriented standards and approaches, like BSC [117], CMM and CMMI [59, 60, 110], PSM [58, 149], CobIT [19], ITIL [20], and ISO standards, advocate the importance of the measurement alignment with business and organizational issues.

The misalignment problem and the existence of a gap between measurement programs at the operational level, e.g. projects, and organizational strategies motivated the development of the GQM*Strategies approach.

### 2.3 GQM*Strategies Approach

The GQM*Strategies [13, 15, 17] is the result of a 30-year-long evolution and use of the GQM method [3, 12, 26, 131, 143, 157, 163, 164, 189, 196]. The foundation of the

---

16Quality Improvement Paradigm [7], also explained in section 2.3.2
17BSC–Balanced Scorecard
18Practical Software Measurement
19Control Objectives for Information and related Technology, http://www.isaca.org/Knowledge-Center/cobit
method was laid out in the Software Engineering Laboratory (SEL) at the beginning of the 1980s [21].

GQM provides a method for an organization or project to define goals, refine those goals down to the specifications of the data to be collected, and then analyze and interpret the resulting data with respect to the original goals. However, GQM does not provide a mechanism for linking high-level business goals to lower-level goals or for supporting and integrating different goals at different levels of the organization. Such a mechanism is provided by the GQM+Strategies.

A major change is that the GQM+Strategies is designed to address the business aspects of an organization in a goal-driven fashion and allow business people to speak in their language—the language of business—when developing measurement programs. The change is logical for an approach that aims to be a wider framework for integrating different goals from different organizational units and levels, and only the language of business is universal enough to accomplish that.

The GQM+Strategies package includes the following two components [17]: the meta-model—the definition of concepts used for modeling real-world situations, and the process (the grid derivation process)—the definition of the steps and tools (assets) for implementing the GQM+Strategies in a real-world setting. In the following sections, each of the components is elaborated.

2.3.1 The Meta-Model

The GQM+Strategies meta-model introduces several new concepts: multi-level goals, strategies, context/assumptions, and an enhanced multi-level interpretation model. Figure 9 depicts the GQM+Strategies concepts and their interrelations. The new concepts are on the left side, and from those concepts the core concepts are goal, strategies and context/assumption. The remaining concepts of the Goal+Strategies element and levels are introduced with the purpose of more efficiently handling GQM+Strategies structures, i.e. grids. The Goal+Strategies element (Figure 9) represents a single goal and its derived strategies, including all context factors (facts about the business environment) and assumptions (predictions) that focus and bound the goal and corresponding strategies.

Discernment is made between an organizational goal and a GQM goal. The former is an objective for which strategies need to be developed to accomplish it. The latter is the associated measurement scheme (metrics and interpretation model).
The organizational goals are further refined with strategies. Strategies in turn generate lower-level organizational goals. Organizational goals are formally documented using the goal template with eight dimensions as shown in Table 3, while strategies are documented in the form of statements. An example strategy [16], $\sigma$, for the organizational goal $G_1$, given in Table 3, is formulated as:

$\sigma$: test reliability in (e.g. remove more defects)

At the same time, the goal ($G_1$) and the strategy ($\sigma$) are accompanied by assumptions about the effects of customer satisfaction on customer loyalty (A1), how to measure...
customer satisfaction (A2), and assumptions about how customer satisfaction can be improved (A3, A4). The environment of the top-level goals is characterized by context elements C1, C2, and C3 (Table 6).

Furthermore, strategy $\sigma$ refines goal G1 into goal G2 at the next lower level. Goal G2 is given in Table 4.

### Table 4. Goal formalization template for an example goal $G_2$.

<table>
<thead>
<tr>
<th>Goal Dimensions</th>
<th>Formalized Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>Decrease</td>
</tr>
<tr>
<td>Focus</td>
<td>Customer reported software defects</td>
</tr>
<tr>
<td>Object</td>
<td>System test process for Splash</td>
</tr>
<tr>
<td>Magnitude</td>
<td>20% reduction in number of customer reported defects</td>
</tr>
<tr>
<td>Timeframe</td>
<td>12 weeks after release</td>
</tr>
<tr>
<td>Scope</td>
<td>Web Products Division, Splash Project Manager</td>
</tr>
<tr>
<td>Constraints</td>
<td>Splash price and functionality</td>
</tr>
<tr>
<td>Relations</td>
<td>Development cost goals, schedule goals</td>
</tr>
</tbody>
</table>

Goal G2 is associated with the fact that the expertise for a new testing system and process were successfully adopted (C4). Further on, goal G2 is addressed with strategy $\sigma'$: *introduce the new system test process*. It is expected that reducing slippage by 20% reduces reported defects by 20% (A5). The strategy $\sigma'$ connects goal G2 with goal G3 at the project level. Goal G3 is given in Table 5.

### Table 5. Goal formalization template for an example goal $G_3$.

<table>
<thead>
<tr>
<th>Goal Dimensions</th>
<th>Formalized Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>Decrease</td>
</tr>
<tr>
<td>Focus</td>
<td>Defect slippage</td>
</tr>
<tr>
<td>Object</td>
<td>New system test process for Splash</td>
</tr>
<tr>
<td>Magnitude</td>
<td>20%</td>
</tr>
<tr>
<td>Timeframe</td>
<td>12 weeks after release</td>
</tr>
<tr>
<td>Scope</td>
<td>Web Products Division, Splash Project Manager</td>
</tr>
<tr>
<td>Constraints</td>
<td>Splash price and functionality</td>
</tr>
<tr>
<td>Relations</td>
<td>Development cost goals, schedule goals</td>
</tr>
</tbody>
</table>
The list of all context and assumption elements for the *Splash product* example is given in Table 6.

**Table 6. Context and assumption elements for the Splash GQM+Strategies grid example [16].**

<table>
<thead>
<tr>
<th>Context/Assumption</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context C1</td>
<td>Highly competitive market for XYZ class of products.</td>
</tr>
<tr>
<td>Assumption A1</td>
<td>Improving customer satisfaction will increase customer loyalty.</td>
</tr>
<tr>
<td>Context C2</td>
<td>Little control over development process.</td>
</tr>
<tr>
<td>Context C3</td>
<td>Limited budget.</td>
</tr>
<tr>
<td>Assumption A2</td>
<td>Satisfaction can be measured by # of complaints.</td>
</tr>
<tr>
<td>Assumption A3</td>
<td>Many complaints are due to product reliability.</td>
</tr>
<tr>
<td>Assumption A4</td>
<td>Reducing defects by 20% reduces complaints by 10%.</td>
</tr>
<tr>
<td>Context C4</td>
<td>New system test process available.</td>
</tr>
<tr>
<td>Assumption A5</td>
<td>Reducing slippage by 20% reduces reported defects by 20%.</td>
</tr>
<tr>
<td>Context C5</td>
<td>Defect slippage data available on past projects.</td>
</tr>
<tr>
<td>Assumption A6</td>
<td>Baseline projects relevant.</td>
</tr>
</tbody>
</table>

The *GQM graph* is a collection of GQM goals that measures a GQM+Strategies element. For example, the goal G3 is measured through the following GQM goal [13]:

**MG3:** Analyze the system test process in order to evaluate it with respect to a 20% improvement in defect slippage compared to past projects from the point of view of the test manager of the Web Product division.

The measurement goal MG3 specifies the Defect Slippage Ratio (DSR) as the ratio between the quotient of faults found in the system test and faults found after the system test on the project and the same quotient in a historical baseline set of projects. If the DSR is greater than or equal to 1.2, then there is at least a 20% improvement.

The *interpretation model* (Figure 9) integrates GQM graphs from different organizational levels and specifies an interpretation logic that is defined by a hierarchy of Goal+Strategies elements. The example given in [13, 16] shows the logic of the GQM+Strategies interpretation model in the form of a flowchart; in other words, the interpretation model’s logic is defined by a set of rules (i.e. if-then rules):
//Example of the GQM+Strategies interpretation model

interpretation_model_start:

IF CCR \leq 0.9 \text{ THEN } //CCR: Customers' Complaints Ratio
    Goal (G1) of 10\% in customer complaints is ACHIEVED;
ELSE IF SDCC \leq 0.8 \text{ THEN } //SDCC: SW Defects Customers Complaints
    Goal (G2) of 20\% improvement in unique customer related SW defects is ACHIEVED;
    Check assumptions connecting G1 and G2;
ELSE IF DSR \geq 1.2 \text{ THEN } //DSR: Defect Slippage Ratio
    Goal (G3) of 20\% improvement in defect slippage is ACHIEVED;
    Check assumptions connecting G2 and G3;
ELSE
    Reconsider strategies (i.e. action plans) for G3;
    Check context and assumptions for G3;
ELSE
    Reconsider strategies for G2;
    Check context and assumptions for G2;
ELSE
    Reconsider strategies for G1;
    Check context and assumptions for G1;

interpretation_model_end.

The interpretation model provides information not only about the goal’s achievement, but also about the effects of strategies and the correctness of the assumptions. In the example above, if goal G2 is not achieved but goal G3 is achieved, the attention will be directed to the connecting assumptions, i.e. the assumption that reducing defect slippage by 20\% will reduce customer reported defects by 20\% might be wrong, and it may be that 30\% or more of defect slippage reduction is needed to achieve goal G2. In a case when data are out of chart readings, then the “fine tuning” of the assumptions is not sufficient, and such situations require serious reconsideration of the strategies.
Measurement goals (GQM goals) can be project-specific goals, dealing with various project-related issues, e.g., improving testing efficiency. At the same time, those measurement goals are a part of a larger structure—a grid—and when interpreted in the context of the grid, they measure and assess the success of the organizational goals and strategies. Having the ability to measure the effectiveness of strategies and to analyze the impacts of strategies on different organizational goals was referred to as organizational reasoning about goals and strategies in chapter 1.

2.3.2 The GQM*Strategies Process: The Grid Derivation Process

The recommended adoption model for the GQM*Strategies approach is the QIP model [8], which is referred to as the GQM*Strategies process (Figure 11).

The QIP paradigm supports continuous process improvement and engineering of the development processes, as well as helping with technology infusion [8]. The QIP paradigm sees software discipline as evolutionary and experimental, and it takes premise that all project environments and products are different. The QIP model is
based on two closed loop cycles, i.e. an organizational (larger) and a project (smaller) cycle. The project-specific feedback cycle provides feedback to the project during the execution phase in order to prevent and solve problems, monitor and support the project, and realign chosen processes with defined goals. In the case of the GQM+Strategies approach, selected strategies can be seen as organizational projects that are executed to achieve goals. The organizational feedback cycle provides feedback to the organization after the completion of the project. The purpose of the organizational feedback is to analyze the concordance and discrepancy of the collected data against previous experiences and models. This helps to increase the understanding of the concluded experience and to capture experience in a form that can be used by other projects.

Fig 11. The QIP model of the GQM+Strategies process.

The organizational cycle represents how an organization learns. Before initiating the GQM+Strategies process, an initial step is performed to define resources and commitment (Figure 11). The process itself is divided into the following six phases:
1. **Characterize and understand.** In this step, the scope of the GQM+Strategies application is defined, and the organizational environment is characterized, i.e. general context elements are defined.

2. **Set goals.** First, an organizational structure is determined and gap analysis is performed, followed by the prioritization of goal and the GQM+Strategies grid derivation for the organizational units and levels within the application scope.

3. **Choose process.** This step is about planning and organizing the implementation of the strategies. Implementation plans are defined, data collection and analysis is organized, and feedback mechanisms are established.

4. **Execute process.** The GQM+Strategies grid is executed, i.e. strategies are applied, data collected and validated, and feedback sessions performed. The execution of strategies is done by different organizational units, and for each of them, it presents a project cycle with project learning mechanism that is divided into three activities:
   
   a) **Process Execution.** This represents implementation of strategies for local goals, i.e. goals specified by the organizational unit.
   
   b) **Analyze Results.** The strategy implementation plans are monitored and analyzed. If needed, some local changes are made. In other words, changes that do not require reviewing upper-level goals and strategies are acceptable.
   
   c) **Provide Process with Feedback.** The knowledge and experience is packaged for sharing with other units and departments within the organization.

5. **Analyze results.** This step provides an opportunity to analyze data and revise strategies if needed. Findings are shared and communicated.

6. **Package and improve.** This step is to adapt and improve the grid, correct wrong assumptions, and set new ones. Strategies should be adapted if assumption changes require that.

A central activity of the QIP phase of setting goals is a grid derivation process. The grid derivation process spreads across organizational units and levels that are within the GQM+Strategies application scope.

**The Grid Derivation Process**

The grid derivation process [16] defines activities and steps that have to be undertaken in order to produce a GQM+Strategies grid, which involves: capturing or identifying organizational goals and strategies, formalizing goals with the goal template, identifying
relationships between goals (e.g. strategies), developing appropriate measurement schemes, and finally defining interpretation model.

The GQM+Strategies grid derivation process [16] is flexible and allows different approaches, starting from top-level business goals down to addressing lower-level goals or vice versa. During the derivation process, two parallel threads are running (Figure 12): (1) one is related to defining business goals, context/assumption elements, and strategies for addressing goals, and (2) the other is related to defining measurable goals and actually deriving the GQM graph.

![Diagram](image.png)

Fig 12. The GQM+Strategies grid derivation process.

In the following an overview of the grid derivation process is given [16].

**Elicit General Context and Assumptions.** The organizational (business) environment is modeled by specifying context factors and assumptions. For example, the corporate vision and mission statement represent general context elements for top-level business goals.
**Define Top-Level Goals.** First, an initial set of high-level goals is identified. Second, the goals have to be prioritized and analyzed for potential conflicts. Third, the selected goals are formalized using the GQM+Strategies goal template (Table 3).

**Make Strategy Decisions.** First, a list of potential strategies for achieving the business goals is identified. Second, the most promising strategy has to be selected, considering the cost and benefit analysis.

**Define Goals.** If it is possible to refine the strategy by another goal level, the implications of the chosen upper-level strategies with respect to lower-level goals first have to be determined. Second, potential lower-level goals are identified based on the analysis. Third, the most promising goals with respect to feasibility, cost, and benefit is selected. Fourth, the most promising goals are selected and formalized using the GQM+Strategies goal template.

Creating the measurement branch of the grid for each goal and strategy level is not an isolated task; that is, the metrics derived across different levels of the GQM+Strategies model will usually overlap. Moreover, an interpretation model for a higher-level goal may be defined completely only if the lower-level pieces have already been modeled.

**Define GQM Graphs.** The GQM Graph derivation process is well-documented in the literature\(^{21}\).

The whole process is iterative and involves people, or representatives, from various organizational levels. In creative industries, such as the software industry, the involvement of people or groups of people at different organizational levels is essential for success. The GQM+Strategies approach involves people in a cooperative and constructive way, by asking them to define their own goals and strategies that will support upper-level goals and strategies. This form of flexibility allows people at different levels to define goals and strategies within their comfort zones that are constrained with upper-level goals and strategies. In other words, people at different levels are provided with **bounded flexibility** that allows them to optimize locally while they are aligning their activities with global organizational goals and strategies.

### 2.4 Value-Based Software Engineering

The VBSE addresses a similar set of problems as the GQM+Strategies approach. Furthermore, the VBSE is well established in software engineering community, and

---

it connects different management theories in the context of software engineering. Therefore, it is a good framework for guiding further improvements and additions to the GQMStrategies approach.

The goal of software engineering is to create products, services, and processes that add value [29, 179]. Software service providers and product developers need to make different kinds of decisions (e.g. technological or architectural) that can impact value creation. If the key decisions remain implicit and hidden that can lead to negative effects on the value creation [29]. The decisions are driven by individual or collective goals and expectations of gaining some benefits (i.e. value realization). According to Biffl, Aurum, Boehm, Erdogmus & Grünbacher [29], the expected benefits can be tangible or intangible, economic or social, monetary or utilitarian, or even aesthetic or ethical. In other words, the value is a multidimensional concept that admits multiple characterizations chosen or set by the value beholder. As such, the concept of value cannot be desubjectified.

In software and IT-enabled businesses, the business value is limited less by hardware capability and more by the ability of managers to invent new processes, procedures, and organizational structures that leverage this capability [53].

The agenda of the VBSE [36, 37, 43] was proposed by Boehm. The main objective of the value-based software engineering is to incorporate the value perspective into the software development process and activities. Applying the VBSE principles establishes a value sensitive environment, and software products and services developed in that environment are engineered to meet value criteria of different stakeholders.

The key principles of the value-based software engineering are [36, 37]: benefits realization analysis, stakeholder value proposition elicitation and reconciliation, business case analysis, continuous risk and opportunity management, value-based monitoring and control, and change as opportunity.

1. The Benefits Realization Analysis. All initiatives needed to realize the potential benefits of a system are identified and coordinated. The Benefits Realization Approach [187] accentuates the importance of techniques for modeling and value assessment of the investment initiatives (programs). For modeling the Result Chain tool is suggested, while for value assessment, a set of four predefined questions is used (Are we doing the right things? Are we doing them the right way? Are we getting them done well? Are we getting the benefits?) [187]. Although the Benefits Realization Approach is mainly used by IT-enabled business, it can be beneficial to business value analysis in general.
Thorp [187] identified four so-called management blind spots when dealing with organizational value: linkage—non-existing or vague links between expected results of a business strategy and investments done on different organizational levels, reach—an unclear picture of the breadth of change required by an investment, people—lack of proper motivation and preparation of the people who will be affected by change, and time—notoriously hard to estimate realistic time frames. Business value analysis should shed light on those blind spots. Linking resources to outcomes increases the concreteness of a software project and helps to identify stakeholders who need to be involved in system development. The analysis also results in visible contributions, outcomes, and assumptions about the system [36, 39].

2. Stakeholder Value Proposition Elicitation and Reconciliation. This involves identifying and documenting success-critical stakeholder value propositions. Approaches that help in agreeing upon the value propositions include expectations management, visualization and tradeoff-analysis techniques, prioritization, and the use of groupware tools [36].

The groupware tools and approaches (e.g. reviews, workshops) are also important for managing knowledge [201]. Therefore, an important process of knowledge building and sharing takes place behind the value proposition elicitation and reconciliation activities. Boehm [39] suggests using the Value Proposition Model-Clash spiderweb diagram as a technique for visualizing value propositions.

3. Business Case Analysis This involves determining relative financial costs, benefits, and return on investment (ROI) of a system during its life cycle. ROI is expressed as:

\[ \text{ROI} = \frac{\text{benefit} - \text{cost}}{\text{cost}} \]

When costs and benefits are not occurring at the same moment in time then the time effects are taken into account by discounting values, e.g. by calculating Present Value [39]. Unquantifiable benefits, such as stakeholder good will and uncertainties, e.g., assuming that a product will be the first of its kind in the market, make the business case analysis challenging. Analyzing uncertainties helps in identifying risks related to each development option [36], and it can be used for analyzing value-of-information for justifying buying information (e.g. purchasing market analysis report) to reduce risk [39]. Furthermore, value analysis offers the following benefits to decision-makers [186]:
– Clear definitions of “value” and how those definitions relate to decisions from all stakeholders’ perspectives;
– Analysis of the tangible, quantitative outcomes of the project, but with credibility testing to avoid unfounded conclusions;
– Clear, logical analysis of non-quantified and intangible factors;
– A compelling vision that provides logic and significance for non-quantified and intangible factors;
– Quantifying the value of flexibility, options, and choices;
– A roadmap that helps firms navigate the transformation from current state to future state; and
– Metrics and analysis to manage projects during and after deployment to capture value, control risks, and capitalize on opportunities.

When properly implemented with supporting information systems, it has the potential to lead to better business decisions and results [186].

4. Continuous Risk and Opportunity Management means that risk analysis and risk management should be carried out during the entire life cycle of the system. Risk management involves understanding and addressing people’s utility functions and using risk to determine how much is enough, e.g., when to stop testing, so that resources are not wasted on inessential points [36, 108].

A key for successful risk and opportunity management is to understand people’s utility functions [39]. Understanding the needs and affinities of individuals increases for the chance of discovering proper resolutions when unexpected events occur.

5. The Concurrent System and Software Engineering stresses using iterative process models instead of waterfall style models. Concurrent process models typically have milestones that provide common commitment points [36]. For example, the eXtreme Programming paradigm is well-aligned with the VBSE approach [103].

6. Value-Based Monitoring and Control. This aspect deals with monitoring the realization of the business value of outcomes at the project and organizational levels. It must be noted that even though a project is very successful with respect to its cost-related value, it could be a disaster from the organizational value viewpoint [36]. Figure 13 illustrates the basic value realization feedback process. Business cases are monitored
for their performance according to plans, and while the value is being realized and no changes in assumptions are detected, there will be no corrective actions defined [39].

Fig 13. Value realization feedback process. Source: [39].

At the organizational level, Boehm [39] suggested the use of the Balanced Scorecard approach [116, 118] that organizes goals, strategies, and initiatives into four perspectives: financial, customer, internal business processes, and learning and growth; or the use of the BTOPP approach [187] that uses five perspectives of business, technology, organization, process, and people.

7. Change as Opportunity means realizing that the ability to adapt to change has business value, as the rate of change is continuously increasing. Companies that can react quickly will be more successful. The philosophy is very similar to that used in agile software development [103].

Since the promotion of the value-based software engineering field in the late 1990s, the field has gathered researchers and practitioners around relevant VBSE topics, such as requirements engineering [84, 133], software architecture [123], design and implementation [175], quality assurance [80, 107, 108, 168], risk management [25], and other various aspects of software development (Figure 2 in section 1.1.1). The concepts of the VBSE, at first, look to be strong advocates of economics and business principles. This is not surprising, because the idea came from an author who is a pioneer in the area of software engineering economics (see [35]). However, the theoretical framework of the value-based software engineering rests on five theories that are important and relevant for engineering practice and for the business aspects of corporate economics, and it involves more "soft" managerial and social theories [45].

72
In the following sections, an overview of the VBSE theoretical framework is presented.

### 2.4.1 The "4+1" Theory of the Value-Based Software Engineering

The theoretical framework for the value-based software engineering rests on five interconnected theories [44–46], of which Theory W is central (Figure 14); furthermore, it was developed specifically to address issues of the software development organizations [46]. The VBSE theory is designed to address considerations of computer science theory, managerial aspects of software engineering, personal and cultural aspects of software organizations, and economic values involved in developing and evolving successful software-intensive systems [44].

Figure 14 depicts components of the VBSE "4+1" theory. At the core of the theory lies the success-critical stakeholder (SCS) win-win Theory W. Theory W addresses the questions: what values are important? and how is success achieved? In order to answer those questions it is important to identify all success-critical stakeholders, which is a task of dependency theory.

The dependency theory ((2) in Figure 14) addresses the question how do dependencies affect the value realization? By analyzing the value realization process, i.e. the sequence of situations and outcomes that lead to the value realization, it is possible to identify key stakeholders in that sequence. For that purpose, the use of the Results Chain tool [187] was suggested as a good solution for visualizing outcomes, contributions, and related assumptions [44, 45]. As a result of understanding dependencies, the new success-critical stakeholders that need to be involved might be identified. The newly identified stakeholders and their expectations and contributions are added to the result chain. Finally, the result chain will involve stakeholders who are critical not only for software-related activities, but also for non-software-related activities, and their joint focus leads to value-producing end results [44].

After identifying all success-critical stakeholders (SCS), it is important to understand how the SCS want to win, as well as what their winning conditions are. This is addressed by utility theory ((3) in Figure 14). The basic question of the utility theory is how important are the values? This is a complex question that involves a whole spectrum of
issues, starting from physiological and social theories like Maslow’s need hierarchy\textsuperscript{22} up to issues of bounded rationality and Simon’s satisficing theory\textsuperscript{23} [45].

Having identified all critical stakeholders and their winning conditions (i.e. value propositions) the next step is to negotiate win-win decisions [41] in a way that all stakeholders are satisfied, in other words, to reach the win-win equilibrium. These issues are addressed by decision theory ((4) in Figure 14). For a successful win-win equilibrium different techniques can be used, such as WinWin negotiation model [44].

The control theory component (Figure 14) addresses issues of controlling progress toward SCS win-win realizations. The principal question that is addressed by control theory is how to adapt to change and control value realization. Organizations need to meet the following conditions of the successful enterprise control [44]: (1) ability to observe the current enterprise state, (2) ability to predict whether the enterprise is heading toward an unacceptable state, (3) ability to redirect the enterprise toward an acceptable near-term state and a successful end state, and (4) ability to avoid positive feedback cycles that cause control systems to overcompensate and become unstable. In the VBSE, the control theory is used with an accent on managing the expected value being realized by a project, and not on just tracking the project’s progress against plans [44].

\textsuperscript{22}\text{see [144]}
\textsuperscript{23}\text{see [181]}

\textbf{Fig 14.} The Overall Structure of the “4+1” Theory of VBSE. Source: [44].
In practice, the five theories of the VBSE interact together. Most of that interaction is happening with or is realized through Theory W [44, 45].

### 2.4.2 Theory W

Software development organizations as knowledge-creating organizations [158] depend on the creativity and innovativeness of the various groups of people across the organization, e.g. developers, product or project managers, and business managers. Setting up an environment where the people can collaborate and work together toward common goals is a concern of management theories and practice.

Different management approaches and styles can have positive or negative impacts on creativity and innovativeness [136, 156]. Boehm & Ross [46] compared and analyzed different organizational management theories on the issues that software organizations are facing. The selected theories (referenced as X, Y, and Z theories in [46]) were: Theory X—Frederick Taylor’s scientific approach, Theory Y—Evans’ theory of productive software management, and Theory Z—Gellerman’s work on motivation and productivity. The result of analysis was that each of those management theories or approaches fall short when addressing adequately the issues of software organizations, and this shortcomings motivated the development of Theory W.

Theory X (Taylor’s work) perceives people as being as predictable as machines, which gives no room for creativity or adaptiveness; it also kills self esteem and leads to an organizational inability to deal with changes [46, p.903]. Theory Y (Evans’ work) perceives managers as coaches with a role to stimulate individual creativity and innovativeness. Furthermore, it accentuates the individualism that leads to agile organizations capable of coping with changes. On the other hand, it creates difficulties in dealing with conflicts [46]. Theory Z (Gellerman’s work) addresses the issues of Theory Y by accentuating the importance of the organizational shared visions and establishing an organizational culture that should minimize potential conflicts and set a proper environment for resolving conflicts when they occur.

The fundamental principle of Theory W is that organizations can be successful only if all actors (stakeholders) are winners, and that the primarily role of the managers is to lead the organization toward situations where all parties will be winners, i.e. a win-win situation or win-win equilibrium.

At the core of Theory W is the enterprise success theorem [44, 46]:

---

75
Your enterprise will succeed
if and only if
it makes winners of your success-critical stockholders.

Boehm & Jain provided an informal proof for the enterprise success theorem [44]. The arguments for the proof of “if” are that everyone significant is a winner and that nobody significant is left to complain, while the arguments for the proof of "only if" are that nobody wants to lose, that prospective losers will refuse to participate, or will counterattack, and that the usual result is lose-lose [44].

While the proof of "if" is clear, the argumentation for "only if" is not so clear and requires further persuasion. Boehm & Jain [44] illustrated the proof of the theorem as an example. They made a point that win-lose situations tend to end as lose-lose situations, assuming that nobody will start from a lose-lose situation in the first place. Lose-lose situations can happen because not all stakeholders practice their shares of interest at the same time. For example, a customer and developer can agree on conditions that make them winners, while a user is left out as the loser (win-lose situation) [44]. Such a project can run smoothly and deliver a product, but when the product needs to be valorized by the market, the unsatisfied users reject the product, and it fails to generate the expected revenue. The win-lose situation ended as lose-lose. Therefore, it is important to identify win-win situations from the beginning. In order to achieve that, it is necessary to understand how people want to win, based on established reasonable expectations that are matched with people’s tasks and their win conditions, as well as to establish a supportive environment for the people [46].

The significance of the “4+1” theoretical framework is that it discusses what and in what way the existing theories and bodies of knowledge are relevant for the software engineering field in order to have a holistic view of the software development activities. The most relevant aspects of those theories are abstracted with the seven key elements of the value-based software engineering framework, which eases the application of the relevant theories into the software engineering context. Such applications include contributions to the areas of value-based software metrics and risk management.

2.4.3 Value-Based Software Metrics

The VBSE metrics are used for supporting the value creation process. In that context the metrics are utilized for a variety of different applications: software process improvement
characterizing software product lines [34], supporting software development activities [134], developing cost–benefit models [168], predicting system properties from its design [175], justifying software reuse [71], etc. However, earned value management (EVM) and earned value metrics and indicators are most commonly used for the purpose of value-based monitoring and control (the VBSE key practice #6) [101, 113, 194].

Earned value management is a quantitative approach for monitoring project performance [50]. The approach is particularly useful for controlling projects’ budgets and schedules, and estimating actual completion times and costs based on current project performance. The origins of the earned value concept date back to the 1950s and 1960s. First, the Program Evaluation Review Technique (PERT) approach was extended to include cost variance and basic concepts of earned value. The approach did not survive, but the idea of earned value did and it was included in a new approach called Cost/Schedule Control Systems Criteria (C/SCSC) [50]. The EVM was used successfully as a technique for managing software development projects [56, 67, 112]. However, Humphrey et al. [112] noticed that the abilities of projects to control cost and schedule variance depend on the maturity of the development processes used.

Brandon [50] defines earned value as the value (usually expressed in some currency) of the work accomplished up to a point in time based upon the planned value for that work. In order to perform an earned value analysis of a project it is necessary to collect basic earned value metrics [43, 50]:

- **Budgeted cost of work schedule (BCWS)**—the total budgeted cost up to the analysis date. The BCWS represents planned costs up to the analysis date, and it is time dependent only.
- **Actual cost of work performed (ACWP)**—this is what it actually cost to accomplish all the work completed as of the analysis date. The ACWP or just cost tells what was spent on the project so far.
- **Budgeted cost of work performed (BCWP)**—the cost originally budgeted to accomplish the work that has been completed as of the analysis date. Additionally, this represents the earned value. If the project is capable of delivering more in the planned time, then an organization will “earn” because it will spend less resources than was originally planned.

One advantage of the earned value system is that it provides a possibility to visualize the metrics with plot diagrams for an easier analysis of trends. Figure 15 illustrates one
earned value system for an example software project that goes through four development phases [43].

The basic earned value metrics are used to calculate the cost variance as:

\[
Cost\ Variance = ACWP - BCWP,
\]

where cost variance represent the difference between actual cost and budgeted cost (red and green lines in Figure 15). Another useful cost metric is the cost performance index or cost efficiency factor:

\[
CPI = \frac{BCWP}{ACWP},
\]

and a CPI of less than 1 indicates budget overruns. In a similar way, the schedule variance is calculated:

\[
Schedule\ Variance[Dollars] = BCWS - BCWP,
\]

where schedule variance represents the difference between the blue (BCWS) and green (BCWP) lines in Figure 15. Often, it is more convenient to express schedule variance in calendar time than in dollars:

\[
Schedule\ Variance[Months] = \frac{Schedule\ Variance[Dollars]}{Panned\ Cost\ per\ Month[Dollars]}.
\]
Schedule performance index or schedule efficiency factor is defined as:

\[ SPI = \frac{BCWP}{BCWS} \]

and an SPI of less than 1 indicates that the project is behind schedule.

For the example in Figure 15, it was planned to do prototyping, analysis, and to deliver plans for the new product in four months. The diagram shows that \( BCWP < BCWS \) indicating that the project is behind the schedule, i.e. \( SPI < 1 \), while at the same time project expenditures are within budget (\( CPI > 1 \)). In this particular situation, falling behind schedule can be connected with providing less resources than originally planned, and therefore the project is spending less than planned. In order to obtain accurate values for the BCWP it is necessary to assess the progress of the task or work. In general, the earned value metrics are sensitive to a continuous assessment of task completion. At the project level, a project manager estimates the percentage of task completeness [43]. Such estimations are even more challenging in software development projects, where the definition of “done” is quite flexible. Even when a task is completed, it is not uncommon to have additional rework for the same task. Therefore, more sophisticated mechanisms are needed for assessing task completion in the software development context. One possibility is to utilize project-level measurement programs for that purpose [56, 67]. For example, the EVM approach can be paired with the GQM paradigm to monitor project goals [67], or with burndown charts of agile projects to generate BCWP metrics [56].

Project-based metrics and the earned value approach do not have sufficient capabilities to address the real value creation for an organization [113]. For example, projects are bound in time, while the value realization can take place after the projects completion. Also, the value is context dependent and sometimes it is hard to express it as a scalar quantity [47]. All this leads to difficulties with defining a generic VBSE measurement framework. Some approaches have addressed this issue by proposing models that integrate software processes with a business value perspective [140], but the issues are still open with respect to how to integrate non-software development process (e.g. sales, marketing) into such models.

Therefore, Boehm & Huang [43] proposed integrating critical stakeholders’ views of value with EVM through the benefits–realization approach [187] and risk/opportunity management practices. The benefits–realization approach enables identification of outcomes and assumptions related to the realization of the outcomes. A sequence of outcomes forms a results chain. The analysis of the results chain provides a basis for
risk/opportunity management. Therefore, the value-based earned value monitoring and control should interpret the earned value together with risk/opportunities and benefits realization.

![Diagram](image)

**Fig 16. Value realization feedback process. Source: [43].**

Figure 16 explains how the earned value metrics can be used for the value realization feedback process. The important question in the value realization feedback process is whether *value is being realized* and that is answered by comparing BCWP and BCWS. If the BCWP is greater than BCWS and greater than actual costs, then the value realization performs according to plans and no corrective actions are needed [43].

Furthermore, benefits realizations are often not visible before the end of the project. Therefore, the value-based versions of the EVM system need to bridge the gap between actual work done in projects and benefits realizations that occur after the projects completion. The financial-like indicators (e.g. costs and benefits expressed with monetary values) are preferred for calculating earned value metrics. However, such financial-like indicators could easily lead to neglecting other aspects of the software development process that are equally valuable, such as employee or customer satisfaction.

Using performance indicators is good for making informed decisions because they are capturing trends over past performance and direct management attention to potential issues when unwanted changes of trend occur [27].

### 2.4.4 Value-Based Risk Management

The software development as engineering discipline is impacted by different uncertainties, such as *uncertainty in time*—some events may occur with insufficient or no time to react upon; *uncertainty in control*—inadequate understanding of and authority to influence
important decisions of the development process; and *uncertainty in information*—
inadequate or inaccurate information on which to base decisions [86]. The risk 
handling and management skills and tools are well-discussed in the software engineering 
community, e.g. IEEE Software special issue [42].

In the context of the VBSE, the risk is defined as the threat of not realizing value 
propositions of stakeholders and is often quantified as the *risk exposure* (RE) [46]:

\[
RE = LP \times LM
\]

where *LP* is the Loss Probability factor that represents the probability of an unsatisfactory 
outcome, and *LM* is the Loss Magnitude factor, which represents the magnitude of the 
loss if the outcome is unsatisfactory. One way to present LM is in terms of participants’ 
utility functions, which measures the degree to which the participants become losers 
rather than winners [46].

The risk management process is generic, and it can be applied to many different 
domains, not only software development projects. However, one of the steps or phases 
in the risk management process is risk identification and quantification [173], and that 
particular step is domain-specific. For example, the SEI[^24] is using the *software risk 
taxonomy* [98]. The taxonomy represents a categorization of the risk sources in software 
development processes and organizations. Using such taxonomy can significantly ease 
the process of identifying and analyzing risks. Once the risks are identified, they need to 
be assessed and quantified. In risk management practice, it is very common to quantify 
risks using different probability distributions for the likelihood of the risk occurrence 
and to characterize risk impact as a discrete or continuous variable [86, 98, 173].

One such comprehensive risk management method designed for software engineering 
is *Riskit* [83, 126, 127]. The method uses the following model to define risks [126]:

\[
\text{Risk} \overset{\text{characterized by}}{\longrightarrow} \text{(Probability and Loss)} \\
\text{Loss} \overset{\text{defined by}}{\longrightarrow} \text{Expectations/Goals} \\
\text{Expectations/Goals} \overset{\text{valued by}}{\longrightarrow} \text{Stakeholder(s)}
\]

where the risk itself is characterized by its *probability* of occurring and the *loss*, e.g. 
loss of utility. The loss itself is defined by the expectations or goals that are valued by 
stakeholders. Without a clear understanding of the goals or stakeholder’s expectations 
the risk analysis would be incomplete [127]. Further on, the Riskit defines a full risk

[^24]: Software Engineering Institute, www.sei.cmu.edu
management process with a set of guidelines and useful templates for implementing the 
process.

Successful risk management involves three elements [86]:

- **Repeatable process** that is visible, transparent, and measurable. In other words, a cer-
tain level of process maturity significantly increases the chances of risk management.
- **Widespread access to adequate knowledge.** Risk management reintroduces the human 
  factor into the economic decision-making [39], and making those decisions requires 
  sharing and accessing different sources of knowledge.
- **Functional behavior.** Behavior deals with human aspects of the software process, in-
  cluding human interactions, motivations and incentives, perceptions and perspectives, 
  communication and consensus, and decision-making and risk tolerance

Going back to the main objective of value-based risk management, i.e. to monitor 
and control the threats to value realization, it is evident that abilities to systematically 
define and measure value are important for risk management practices [126].

The following related work is used for designing the research approach of this 
dissertation.

### 2.5 Design Science

Many authors in the field of design science were inspired by the work Herbert A. 
that human made artifacts—technology—are artificial, i.e. non-existent in the physical 
world (environment) prior to its creation. But, when technology is created, it interacts 
with the environment, and becomes a part of it. Therefore, the concern of engineers 
and design science is twofold: building knowledge about the creation process itself, 
and understanding the impacts of the existing technologies on the physical world. The 
latter, could use similar methods and approaches as those used in natural or behavioral 
sciences, e.g. observation and experimentation. For the former, however, engineering 
methods and practices are applied.

In the following section the basic ideas of Simon’s work is presented, without going 
into detail. The various details of the design science are addressed by Hevner *et al.* and 
Wieringa, and their views are discussed in later sections.

\(^{25}\)Herbert A. Simon (1916–2001) winner of the 1978 Nobel Prize in Economics, artificial intelligence expert
2.5.1 Work of Herbert A. Simon

In the late 1960s, Simon gave a series of inspiring lectures on the design of the artificial. The lecture notes evolved into the first edition of the book The Science of the Artificial, published in 1970 [180]. In the book, Simon discusses the anatomy and characteristics of the design process that creates artifacts and solutions.

Simon views the process of creating, or designing, a solution as a search process that interacts with the real world and uses an existing knowledge base (theory). The real world is so complex that it is not possible for a designer to have all the details and information needed at once while developing a solution for a practical problem. Making decisions with restricted information is also known as bounded rationality [180]. The restriction of information can arise from several causes, the most common of which are: the asymmetry of information or making assumptions to simplify complex business environments.

Designing a solution, i.e. searching a solution space, requires considering alternative solutions and choosing the best possible. This is what Simon refers to as a heuristic search. In order to choose the right alternative, different approaches can be used, like factorization and means-end analysis [180]. The design search process is very similar to an optimization problem, where some command variables (i.e. independent variables) are given for a utility function, along with a set of constraints. The goal is to search for optimum values for the command variables. At the same time, however, Simon argues that designers are not optimizing but rather sub-optimizing or, as he calls it, satisficing. The satisficing is a consequence of the bounded rationality.

Three decades after the book was published, a group of researchers in the information systems field developed a research framework that is based on Simon’s ideas [96, 106, 142, 167, 170, 198].

2.5.2 Hevner et al.'s Research Framework

Hevner, March, Park & Ram [106] suggested a research framework for information systems studies (IS) that merges design science and behavioral science paradigms.

Figure 17 represents the framework for understanding, executing, and evaluating IS research by combining behavior-science and design-science paradigms [106].
Research problems are identified by the environment (Figure 17), which comprises people, organizations, and technology. The needs that are later formulated as problems come from individuals or groups as a result of all possible forms of interactions among people, organizations, and technology. Recently, Sein et al. [178] argued that the interaction of researchers with the environment can result in the identification of a valid research problem. Therefore, the environment includes the research setting, e.g., the research project or specific technologies that are the legacy of a researcher’s group.

Wieringa [198] differentiates between practical and knowledge problems. A practical problem is defined as one that deals with the difference between the way the world is experienced by stakeholders and the way they would like it to be, while the knowledge problem is one that deals with the difference between stakeholders’ current knowledge about the world and what they would like to know. A solution to a practical problem can change the world, while an answer to a knowledge problem changes the knowledge about the world. In design science research, those two problem types are recurring, i.e., practical problems lead to knowledge problems, while some knowledge problems can lead to new practical problems, and so on [198].

Identifying relevant practical problems is critical for the proper start of the solution development cycle [106]. The following strategies for investigating problems can be used [198]:

---

26 Problems as perceived by researchers.
Problem-driven investigation involves interaction with stakeholders (practitioners) in order to diagnose problems.

Goal-driven investigation is used in situations where there are no real problems, but, for some reason, stakeholders agree to change a solution.

Solution-driven investigation is typical when existing technology searches for new problem areas to which the technology can be exported.

Impact-driven investigation or evaluative research focuses on analyzing technology impacts through past uses and experiences.

All these strategies are not mutually exclusive; it is possible to use a mixed strategy which will have a consequence on the problem definition by emphasizing certain aspects of the problem.

After the identification of research problems, the next step is to build or develop solutions. The solutions are a result of research activities (Figure 17). According to Hevner et al. [106], these activities alternate between design science (build and evaluate) and behavioral science (develop and justify). Behavioral science and design science have different goals or objectives, but they are mutually supportive. The goal of behavior science is the "truth" [106], specifically the truth in a given context or situation that is used as a norm of the "truth" for building theories [33, 193], while the goal of design science is the utility of designed artifacts [106, 180]. Consequently, theory helps to produce better designs, while utility informs theory. This bidirectional relation is a cycle of assessment and refinement (Figure 17). Assessment through justification and evaluation identifies weaknesses and strengths of theories or artifacts, and then provides inputs for a further refinement of the theories or artifacts.

The knowledge base provides building blocks for constructing solutions (theories and/or artifacts). Hevner et al. [106] differentiates between foundations and methodologies. The foundations are inputs for develop/build activities; for example they include theories, frameworks, instruments, constructs, models, methods, and instantiations. Methodologies provide guidelines for justify/evaluate activities, and they can include data analysis techniques, formalism, measures, and validation criteria. Rigor is achieved by applying foundations and methodologies correctly. Having a well-constructed solution does not imply its usefulness. Therefore, in order to meet the relevance criteria it is necessary to select problems from the environment that correspond to real needs, and afterwards to demonstrate the solution application in the appropriate environment.
Hevner et al.’s framework identifies two types of contributions. One type of contribution is the application of a solution in the appropriate environment. The second type of contribution is an addition to the knowledge base. Contributions to the knowledge base can have different forms, but all of them should carry a scientific merit.

For example, an instantiation of an existing theory in a new domain is a valid contribution. In general, the results of the design research, i.e. solutions, have to comply with a non-routine design criteria, and only such results have a scientific merit [106, p. 81]. In other words, a routine design is follows recipes on how to utilize the existing knowledge base to solve emergent problems. For example, a solution for a web-based exchange of information between a manufacture and dealers is a matter of following good software engineering practices available today to utilize existing technology. The opposite of that situation is the non-routine design, which requires a heuristic approach for selecting appropriate building blocks from the knowledge base.

A good solution that exceeds initial setting, that is first classified as a non-routine design, will became the recipe for similar future solutions developed as routine designs.

In Table 7, seven guidelines of the Hevner et al. research framework are given. The motivation for providing guidelines was to make the whole research framework easier to apply in practice. However, according to Hevner the guidelines are not meant to be followed explicitly, but rather the researchers should decide how to implement and tailor the guidelines in their research designs.

2.6 Empirical Software Engineering Methods

In software engineering, the so-called "advocacy research" approach was often used in the 1980s and 1990s [20, 78, 161, 162]. Fenton et al. [78, p. 87] give an illustrative scenario that describes the approach:

"Authors describe a new concept in considerable detail; recommend the concept to be transferred to practice. Time passes, and other researchers derive similar conclusions. Eventually the consensus among researchers is that the concept has clear benefits. Yet practitioners often seem unenthused. Researchers, satisfied that their communal analysis is correct, become frustrated. Heated discussion and finger-pointing ensues."

This was an explicit massage from Hevner during his lecture on Design Science.
Table 7. Design Science Research Guidelines. Source: [106].

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Guideline 1: Design as an Artifact</strong></td>
<td>Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.</td>
</tr>
<tr>
<td><strong>Guideline 2: Problem Relevance</strong></td>
<td>The objective of design-science research is to develop technology-based solutions to important and relevant business problems.</td>
</tr>
<tr>
<td><strong>Guideline 3: Design Evaluation</strong></td>
<td>The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.</td>
</tr>
<tr>
<td><strong>Guideline 4: Research Contributions</strong></td>
<td>Effective design-science research must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.</td>
</tr>
<tr>
<td><strong>Guideline 5: Research Rigor</strong></td>
<td>Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.</td>
</tr>
<tr>
<td><strong>Guideline 6: Design as a Search Process</strong></td>
<td>The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment.</td>
</tr>
<tr>
<td><strong>Guideline 7: Communication of Research</strong></td>
<td>Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences.</td>
</tr>
</tbody>
</table>

The given scenario is lacking empirical evidences that a proposed new concept is beneficial, such empirical evidence can dramatically change the scenario. The personal and subjective judgment regarding the proposed concept shall be substituted with an objective reasoning based on the empirical evidence. This is the main goal of empirical software engineering [20, 161]. The role of the researcher is to understand the nature of the processes used in software engineering and the characteristics of products created by those processes, analyze relationships between them, and formulate hypotheses that are tested or supported by empirical evidence [2]. This process contributes to and evolves the software engineering body of knowledge. It sounds logical and simple, but the reality of software engineering poses many challenges.

Software products are not produced, but rather they are developed by creative individuals [11]. Such processes are not normally repeated over and over in time, and if they are repeated, they have a significantly high variance that is induced by the human
factor [102]. Secondly, the observation and data collection have to take place in real settings, i.e. in organizations that are developing software, and such settings are not isolated environments. Often, controlling or even considering contextual variables is a challenge in itself [11]. Therefore, a realistic approach would be to start with informal studies, such as feasibility studies that can provide valuable insights regarding the factors involved and, over time evolve those studies with the use of formal research methods. In other words, learn through application and as the knowledge about the phenomenon of the study matures it will qualify for more rigorous ways of empirical validation [11].

Empirical software engineering uses different qualitative and quantitative methods, such as experimentation [115, 199], surveying and interviewing techniques [176]. In the following sections, the methods that are used in this dissertation are presented.

2.6.1 Participant Observation

This method is beneficial for collecting qualitative data regarding behaviors and interactions that are hard to capture and would remain unnoticed by other techniques [64, 177]. Participant observation can be used in different stages of research and in different forms, from a general observation technique to a technique where predefined codes are used to capture the specific aspects of behavior and interactions. There are different classifications and definitions of the participant observation techniques, but the main difference is made between techniques where the observer is actually engaged with the activity being observed, and those when the observer is simply collecting data with the consent of those who are observed, the so-called direct observation [177].

2.6.2 Interviews

Interviewing is another very popular method used by researchers in different fields for collecting qualitative data [177]. The structure of the interview questions can be fully defined beforehand, i.e. structured interviews, or it can have no questions at all, which are called unstructured interviews. Structured interviews are capable of investigating specific issues and narrow topics of interest. The process of analyzing and interpreting data obtained with structured interviews is straightforward [169]. Unstructured interviews are excellent for exploring "unknown territories," when very little is known about the subject of the investigation. However, the data obtained in this way contain a lot of overhead, and analyzing the data consumes a great deal of effort
from researchers. Combining the strengths of both types resulted in the semi-structured type. Semi-structured interviews have predefined questions to address topics of interest, but each question also provides flexibility to guide the interview in different directions [64, 169, 177].

2.6.3 Surveying

The surveying method is one of the most popular approaches in qualitative studies because of its cost effectiveness [64]. Unlike interviews, where the richness of data causes labor intensive analysis, surveys use structured questions that significantly ease the process of data analysis. They are particularly good for collecting empirical data from large population samples.

Depending on the scales used for question answers, different quantitative methods can be used. If nominal and ordinal scales are used, then the descriptive statistic and non-parametric statistical tests can be considered and, if ratio or interval scales are used then the whole spectrum of statistical tests can be used [199].

2.6.4 Coding Techniques

Several techniques are available for analyzing qualitative data that involve different types of coding [151, 177]. The analysis of qualitative data is very costly and labor intensive, and often it is a price that researchers pay for gaining highly valid results. Such a high validity is needed when the purpose of the data analysis is to generate theory, e.g., the grounded theory approach [64]. The constant comparison method is a qualitative analysis approach that is used for constructing theories, building hypotheses, and for "making sense" of qualitative data [177].

The constant comparison method involves several coding techniques of interview transcripts. First, the transcript texts are open coded, or the texts are broken into smaller chunks that can be associated with the topics of interests. The open codes can be preformed or postformed [177]. During the coding process, where necessary, the postformed codes or sub-codes are defined to better fit the data, rather than to interpret the data in order to apply some of the preformed codes.

After the open coding, the coded texts are examined for underlying themes or explanations in a process called axial coding [177]. The axial coding is a form of synthesizing propositions or findings that are supported by qualitative data. This process.
is iterative, i.e. formulated propositions are constantly compared with qualitative data for supporting or contradicting evidences. After each iteration, the propositions are reformulated to better represent qualitative data. During each iteration, the data are analyzed and interpreted in a way that they "make sense," this sense-making activity is called selective coding [177]. Selective coding intertwines with axial coding, and it is hard to separate them during the analysis process.
3 Research Approach

Essentially, all models are wrong, but some are useful.
by George E. Box

The name software engineering was coined at the NATO’s First Conference on Software Engineering in 1968, with the purpose of addressing emergent issues with software "production." Since then, researchers and practitioners have reached an agreement in defining software engineering as a disciplined and systematic application of engineering methods to the real-world problems. Shaw [179] gives her definition of software engineering and concludes that all definitions of engineering have common points or goals, i.e. creating cost-effective solutions, dealing with practical problems, the application of scientific knowledge, building things, and in the service of mankind [179, p. 15], which are common understandings of what any engineering is, including software engineering. Consequently, research in software engineering needs to provide a support for practitioners in achieving those goals.

Glass, Vessey & Ramesh examined the state of practice in computer science research [171] and software engineering research [89] and found that the most common research approach, in both fields, is the formulative approach with the use of conceptual analysis as a research method. Interestingly, the same study showed that evaluative research is much less present than formulative (around 10% evaluative studies vs. 80% formulative studies in computer science, and 14% vs. 55% in software engineering research), which suggests that the research generates contributions that are not validated in any kind of empirical studies.

For the advancement of the software engineering body of knowledge, an empirical feedback for proposed technologies (e.g., processes, methods, tools, etc.) is required [19]. The feedback informs theory about the usefulness of proposed technologies and problems with the application of technologies in real world settings. Such feedback allows for better understanding of the real world, which is one of the goals of research.

---

28 North Atlantic Treaty Organization
29 IEEE Standard Computer Dictionary: [...] (1) The application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software; that is, the application of engineering to software. (2) The study of approaches as in (1).
30 [...] disciplined application of scientific knowledge to resolve conflicting constraints and requirements for problems of immediate, practical significance.
work. Any knowledge creation activity tries to generate a new insight or information about the **real world**. The complexity of the real world is beyond human comprehension. Therefore, researchers use simplified representations of the real world, or **models**. The models represent narrow views, or abstractions, of the real world that discard all details that model creators consider irrelevant for a particular view. Traditionally, the modeling of the real world is an important tool that scientists use to develop theories about real world phenomena. In software engineering, modeling is used not only by researchers, but also by practitioners.

Researchers and practitioners in the field of software engineering (referred to as engineers) and researchers in traditional sciences (to differentiate from scientists and researchers in engineering disciplines, they will be referred to simply as **scientists**) are knowledge workers. They apply knowledge, learn from experiences, and contribute new knowledge to a body of knowledge. However, there is a difference in the nature of the work done by those two communities. Blum [33, p. 26] shows through the analysis of relevant philosophical views that scientists are **uncovering** hidden causal relations that are defined by the laws of nature and developing **models** of the real-world to describe the discovered causal relations. Furthermore, if such models are valid, they have predictive capabilities, i.e., the ability to predict and explain the behavior of natural phenomena under study. In contrast, engineers develop **concepts**\(^{31}\) for describing real-world phenomena with the purpose of **solving** real-world problems. Well-developed concepts have a different kind of predictive capability. The predictive capabilities in engineering are in a form that—for certain problems or types of problems, under given circumstances—a solution or type of solution will or will not be effective. The problems are mainly generated by the needs of individuals, groups, and organizations, or as a result of the interactions of individuals, groups, and organizations. This is the foundation of engineering science and **design science** [89, 95, 106, 170, 171].

Blum also noticed that even scientific knowledge is context dependent, i.e., is valid only in certain situations [33, p. 90]. Thus, the main goal of scientists is to seek "universal truths" by decontextualizing knowledge and producing generalizations that are valid in **all** contexts [33, p. 110]. The opposite of that situation is what Blum calls an open-world. The open-world represents plural realities with knowledge assumed to be valid only within certain contexts, and here generalization of knowledge defines a new, larger context in which the generalization is valid. Blum concludes that software

\(^{31}\)A linguistic difference is intentionally made by using the word **concept** instead of **model**.
engineering resides in an open-world and, therefore, the software engineering body of knowledge is contextualized, i.e. context sensitive. Thus, the goal of researchers in software engineering should be exploring and understanding contexts in order to be able to predict the applicability of solutions in different contexts.

Researchers in software engineering need to understand the processes that are used for developing products, and the relationships between process characteristics and product characteristics [2, 10], and how those relationships hold in different contexts. Understanding all context-relevant variables requires different applications by different people with different skills in different settings. Studies of such magnitude exceed the capacity of small research groups and require a network of researchers, which is why Basili [10] refers to this as a big science. For achieving their mission, researchers in software engineering have at their disposal different methods and approaches for developing solutions and for evaluating the solutions together with the industry (practitioners). To properly address complex research problems that have multiple perspectives (e.g., individual, organizational, methodological), it is necessary to employ multiple research methods [152]. Therefore, a research framework is needed to synchronize the use of different methods and to validate the aggregation of the results.

The research framework that is used in this dissertation is inspired by empirical software engineering, in general, and by ideas of design science as discussed above and in section 2.5. The research framework (Figure 18) consists of three main components or three main pillars. Those components are: theory, practice, and the research process.

The components of the research framework that were adopted for this research is explained in the following sections.

3.1 Research Setting: the Role of Theory and Practice

The theory, as a component of the research framework, represents a body of knowledge, a field of related contributions made by other authors. The practice component (Figure 18) represents the accessible real world environments for researchers. Unfortunately, such environments are usually constrained by research practices, which can set a variety of constraints from time and effort constraints, to constraints on research topics, or constraints on organizational involvement with research projects.

The existing theory and research projects are a part of a research setting. Typically, some research setting factors might be predefined, for example, when a researcher
Research projects and work with industry and domain experts cannot influence the decision-making process of the research project preparation phase. These are examples of such factors:

- **Research areas** can be predetermined by the interests and previous experiences of a research group where the research process is taking place. Those interests impact the selection of research topics.
– **Legacy or underlying technologies** can be influential factor. Sometimes research groups (or individuals) might have a history of researching certain technologies or using certain methodologies for which they have established themselves as experts in a scientific community. Those technologies impact the solution development.

– **Industrial partners** are important for providing the real world test environments where newly developed solutions can be evaluated. However, the researchers are bound by contractual agreements between the industrial partners and the research project. Usually, contracts limit information disclosure and access to the resources of the industrial partners.

Those factors have an impact on the quality of research; if handled properly, they can have a positive impact. For example, the familiarity and experience of a research group with certain topics or technologies results in better research plans to address profound research questions.

**Commonalities and Differences with Related Work** The design science research framework (section 2.5) positions the research process between the real world environment and theory, accentuating that research problems should not be solely identified in theory, but also motivated by the business needs of organizations (practitioners). Similarly, Gorschek *et al.* [91] proposed a research approach that involves both academia and industry in order to ensure the practical applicability of the developed solutions. The interaction of theory and practice through the research process in software engineering is possible only if the software phenomenon is studied in its natural environment, i.e. companies and organizations where the software is developed [10].

### 3.2 Research Process

The research process can be divided into two phases (Figure 18). The first phase is *defining the research problem*, which involves not just a review of relevant literature, but also empirical studies with the industry in order to specify more precisely a research problem. When a research problem is formulated, the second phase can commence, the phase of *developing a solution* for a previously formulated research problem.
3.2.1 Phase I: Defining the Research Problem

The research process is initiated by identifying research opportunities. In the case of this dissertation, the direction of the research was predefined by a research project; in other words, the research project provided a selection of possible topics. The topics were proposed by industry partners in the research project. All the topics were of high relevance for practitioners (i.e. industry partners), but they were not all suitable for PhD research. Therefore, the proposed topics are investigated to determine which perceived as valid research opportunities by other researchers in the field. Combining those two inputs made it possible to specify research opportunities within research problem areas (gaps in theory) that are relevant for the industry partners (practice).

The next step is to identify theoretical background and related work, which represents a literature review of previously specified research problem areas. In this step, the related work is identified and discussed, which includes the underlying methodologies, research frameworks, and approaches; i.e., theoretical constructs that are relevant for dissertation work.

After reviewing the existing theories, empirical investigations are needed in order to identify practical problems with existing theoretical constructs. For example, the existing theoretical construct is a methodology that will be further developed in this dissertation. The empirical investigations may include industry pilots (technology transfer), observational studies, or experiments, all of which may use different quantitative and/or qualitative techniques to collect and analyze data from empirical studies. Also, the personal involvement of researchers in empirical studies generates a genuine experience with the objects and methods of the studies, and such experience can help researchers to better understand real world problems.

The formulation of research questions is oriented toward practical problems (i.e. addressing the needs of practitioners). Therefore, the findings of empirical studies are the main input for defining research questions. From a practitioner’s perspective, a solution that successfully solves practical problems has a practical value or practical benefits, while from a researcher’s perspective, such solutions and research work have a high practical relevance.

All the previous steps can be seen as the motivation for a research problem. The research problem is motivated by needs of practice (practical problems) through research questions. However, formulation of a research problem is more oriented toward theoretical challenges and existing gaps in a body of knowledge. In other words, a
familiarity with background theory and related work is needed to fully understand a research problem. Furthermore, developed solutions represent contributions to theory if they are relevant to research problems.

**Commonalities and Differences with Related Work**  All the steps of the problem definition phase are conducted in order to ensure the problem’s relevance for industry practitioners and for the research community. By ensuring the problem’s relevance, the research framework complies with the design science research framework guideline #2 (Table 7), which addresses the business problem relevance. Empirical investigations for identifying practical problems represent Wieringa’s problem-driven strategy for investigating problems (section 2.5.2). Results of the investigations lead to research questions or, as Wieringa calls them, practical problems. Another type of problem is in the form of knowledge problems [198] or gaps in the existing body of knowledge. Those gaps are identified with a literature review of identified problem areas.

### 3.2.2 Phase II: Developing the Solution

After the research problem is formulated, the second phase of the research starts with defining design hypotheses (Figure 18). The design hypotheses are the most important building blocks of the future solutions. They represent constructs that provide a theoretical background for solution development. Solution development is a process with alternatives and options presented in various phases of that process. In order to make that process transparent and criticizable, it is necessary to make explicit the most important decisions that impact solution design and creation. The design hypotheses capture those design decisions, and limit and select solution alternatives.

The design hypotheses also break down a research problem into very specific research subproblems. Note that research questions are used in the process of synthesizing a research problem, while hypotheses are used in the process of analyzing and breaking down the research problem. The former process is driven by practical issues, while the latter process is driven by analyzing theoretical issues and how they can be divided into subproblems. Solving subproblems requires, first, a solution development and, second, the evaluation of hypotheses. Because the design hypotheses cannot be evaluated without solutions, a series of activities is generated in order to develop solutions that use design hypotheses as the underlying theoretical concepts. There are two kinds of
activities, one relating to the solution development and the other relating to the solution validation.

The solution development may involve different techniques, such as conceptual analysis, mathematical modeling, and formal specification. Those techniques are pulled from theory when the ongoing activity requires it. In other words, for the successful completion of research activities, additional literature reviews and queries of the existing body of knowledge might be necessary.

Solution development is followed by solution validation. The goal of solution validation is to provide support for the goodness of the developed solution. The validation is primarily concerned with the correct use of concepts for developing a solution.

The process of defining design hypotheses and activities is an iterative process. It is unlikely that all hypotheses can be defined up front. The more realistic scenario is that, during a solution development for a hypothesis, other hypotheses will be discovered, which would lead to new activities. As the number of hypotheses increases for a particular research problem, a body of knowledge for that research problem will also increase.

Only design hypotheses "tested" thorough solution development are valid contributions to a body of knowledge. Therefore, the needed evidences for supporting design hypotheses are collected during solution evaluations. Note that the previous solution validations were concerned with how well a solution is built, while the evaluations are concerned with how well a solution satisfies its purpose. Thus, the evaluations involve solution pilots in real world settings, i.e. in test environments with collected evaluation feedbacks.

**Commonalities and Differences with Related Work** The solution development phase is compliant with the following design science research guidelines (Table 7). Guideline #1, design as an artifact, occurs in this framework as two types of artifacts result from the research process: the solution itself and theoretical constructs. Here, the design hypotheses specify how different concepts are combined and used; in other words, how new theoretical constructs are proposed. Design evaluation (guideline #3) is conducted during the solution evaluations, while the research rigor (guideline #5) is addressed with solution validations. Simon [180] explains that a design (constructive research) is a search process, a process that searches the solution space in order to find an "optimal" solution. Researchers can use different approaches to search the solution
space, e.g. heuristic [180]. This is the reasoning behind guideline #6, the design as a search process. However, the design science research guidelines are not specific regarding how to conduct the search process nor how to document it. Specifying design hypotheses and activities for developing a solution makes explicit the most important decision points and steps in that search process. For example, the iterative process of defining hypotheses and activities can be seen as a heuristic search process.

3.2.3 Contributions to Theory and Solution Usefulness

The purpose of doing research is to deliver new results. The results of the research process have to meet two criteria—the criterion of a valid theoretical contribution and the criterion of practical usefulness.

Figure 19 shows three layers of the research framework. Each layer represents a different level of abstraction. The bottom layer is concerned with practical problems and deals with concrete industrial cases in research projects, while the upper levels are abstracting practical problems and solutions through a generalization process; in other words, they are building a theoretical model.

The logic of the research framework is the following. Existing needs or practical problems are identified, and they are formulated as research questions. The intention is to define research questions that have wider impact, i.e. that are of interest for other practitioners who are dealing with similar issues. The research questions together with identified gaps in theory help to synthesize a research problem. The research problem is like a long-term research goal or a particular body of knowledge to which researchers contribute. Further on, a research problem generates design hypotheses, i.e. theoretical subproblems that are used for developing solutions. Finally, those solutions are implemented in real world settings.

The ability of solutions to address practical problems represent benefits that the solutions bring (Figure 19). The benefits are observable and measurable. Note that practical problems are expressed by practitioners; therefore they are real and observable, while solutions are the tangible outputs of the research process, and as such they are measurable. The explanations of why a particular solution satisfies a practical problem, and to what degree the solution satisfies it, are provided by the theoretical model behind the solution, i.e. design hypotheses and argumentation explaining how research questions are answered.
The design hypotheses together with research activities are contributions to theory because they create different concepts and explain how the concepts can be blended to address practical problems. In contrast, the applications of solutions in real world environments (practice) demonstrate the practical benefits or usefulness of the solutions (Figure 18).

The logic of the research framework supports the following inference: a solution or a class of solutions will be effective for a problem or a class of problems. There are two main validity issues related to the inference:

- **Internal validity** is about how well a particular solution satisfies practical problems in a given context, and whether the practical problems are relevant.
- **External validity** is concerned with the portability of solutions to different contexts, other than the research context. In other words, external validity discusses the effectiveness of the solutions in different cases and solution limitations.

These validity issues address the top-level inference. However, in various stages of the research process, different research methods can be used, and those research methods have to address their validity threats.

**Commonalities and Differences with Related Work** The design science research framework differentiates two types of contributions: research contributions (guideline #4) and applications to real world environments. In this research framework, the implementations of solutions represent applications to real world environments, while design hypotheses as theoretical constructs are additions to the knowledge base, i.e. research contributions. Extending the knowledge base is done by publishing results.
in peer reviewed conferences and journals (guideline #7). One particularly important attribute of an artifact is its usefulness or utility [95, 106, 170]. The benefits of a developed solution are measurable or observable and they indicate the usefulness of the solution. Practical problems and the application of solutions are what Wieringa [198] identifies as practical problems, while the theoretical model that provides explanations for the research questions is also able to provide answers to Wieringa’s knowledge problems. The internal and external validity issues are suggested in a similar form by Wieringa [198].

3.3 Used Research Methods

This section will give an overview of the research methods used in the dissertation. Table 8 gives a list of methods with references to the points where they are introduced and used in the dissertation, as well as with references to the research framework.

Table 8. Overview of the used research methods.

<table>
<thead>
<tr>
<th>Research Method</th>
<th>Reference in dissertation</th>
<th>Reference to the research framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature review</td>
<td>Chapter 1, Chapter 2, Section 8.3</td>
<td>Research problem areas and opportunities, Identifying theoretical background and related work, Mathematical bases of causality theory</td>
</tr>
<tr>
<td>Analytical paradigm</td>
<td>Section 6.5, Chapter 6, Chapter 7, Chapter 8, Section 8.4.1</td>
<td>Validating solution by conceptual analysis, Developing solution, Validating solution by example</td>
</tr>
<tr>
<td>Interview</td>
<td>Section 2.6.2, Section 4.3, Appendix 1</td>
<td>Collecting pre-study feedback</td>
</tr>
<tr>
<td>Participant Observation</td>
<td>Section 2.6.2, Section 8.4.2, Appendix 4</td>
<td>Collecting evaluation feedback</td>
</tr>
<tr>
<td>Survey</td>
<td>Section 2.6.1, Section 4.5</td>
<td>Pre-study pilots</td>
</tr>
<tr>
<td>Coding Techniques</td>
<td>Section 2.6.3, Section 8.4.2, Appendix 3</td>
<td>Evaluation feedback</td>
</tr>
<tr>
<td>Workshop</td>
<td>Section 2.6.4, Section 4.4</td>
<td>Qualitative analysis of the pre-study feedback data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>101</td>
</tr>
</tbody>
</table>
The literature review as the method of consulting the body of knowledge is used on several occasions. In the beginning, in the phase of defining the research problem, it is used for reviewing the related work and identifying gaps in the literature. Afterwards, during the phase of developing solution, the literature is consulted and revisited for referencing concepts that are used for the solution development.

In engineering disciplines, the analytical paradigm is unavoidable when new concepts and theories are developed. The analytical paradigm is used for solution development and evaluation (the build/evaluate cycle in the design science research framework).

The empirical work with the industry is done through a series of workshops. The empirical data are collected using more qualitative methods like interviews, surveys, and participant observation in combination with quantitative analysis techniques.
II Defining the Research Problem
The Pre-Study: Empirical Investigations of the GQM*Strategies approach

*Education should be exercise; it has become massage.*

by Martin H. Fischer

The aim of these empirical investigations is to study the deployment and use of the GQM*Strategies approach in a real world setting. In particular, the investigation sought out practical problems that practitioners were addressing or trying to solve with the GQM*Strategies grid development. Also, the method pilots were used to collect feedback from practitioners regarding the method’s usefulness and suggestions for further method development. The feedback helped formulate and motivate the research questions.

The empirical research for this dissertation was conducted in two research projects during the period from 2008 through 2011. That time period, presented an opportunity for several pilot projects on the GQM*Strategies approach to be run, mainly with Finnish ICT companies, as well as a chance to interview the practitioners on several of those projects. An opportunity to interview the practitioners from projects run by others also arose. The common goal of all pilot projects was to test the new technology, i.e. the GQM*Strategies approach and later some extensions of the approach. However, the novelty of the approach posed a common challenge for this kind of empirical investigation. The challenge was in transferring the knowledge about the approach to practitioners, because it is possible to investigate the effects and usefulness of the approach in real world settings only if the concepts of the approach are correctly understood by practitioners.

This chapter present only the two pilot projects that allowed for interviews about the effectiveness of the methods for their environments to be conducted. To address the challenge of knowledge transfer, data from the two pilot projects are compared; one, where I was personally involved with technology transfer and one where I was not. Such a cross-case data analysis minimizes the impact of the technology transfer on the pre-study findings.
The data set obtained from the pre-study pilots is rather small and insufficient for a theory generation, i.e. hypotheses testing, but it is sufficient for identifying practical problems and motivating research questions.

4.1 The Pre-Study Design

This phase of the research process used the method pilots as a pre-study for eliciting the needs of practitioners and for evaluating the method’s ability to support practitioners with implementing alignment, communication, measurement, and reasoning cycle (ACMR cycle, Figure 1 in section 1).

The pre-study pilots involved joint work with practitioners on developing GQM+Strategies grids for their goals and strategies. Such joint work included knowledge transfers, brainstorming meetings about goals and strategies, formalizing goals, building grids, and defining measurement schemes for the goals. In other words, it involved a technology transfer process and providing practitioners support in using the methodology.

The research methodology for evaluating the GQM+Strategies approach included the following:

1. **Technology transfer** was conducted with the purpose of transferring basic knowledge and skills regarding the GQM+Strategies concepts and grid development process to practitioners.
2. **Data collection** involved interviewing and participant observation techniques for collecting qualitative feedback from pre-study pilots.
3. **Data analysis** of qualitative data (non-numerical data) was based on the constant comparison method. The method uses several coding techniques, first an open coding for analyzing transcripts, and afterwards axial and selective coding to synthesize findings or propositions.

The interview data analysis results for the GQM+Strategies pre-study evaluations are given in a form of findings in section 4.4 while observations from the pilots that I was involved with are given as reflections and observations (section 4.5).

4.1.1 Technology Transfer

The purpose of the technology transfer is twofold: first, to teach practitioners the method concepts and grid development steps, and second, to provide the practitioners with
a familiar exercise to work with the method. Going through a selected problem area within their own domain helps practitioners to quickly gain the adequate skills and knowledge to use the method. To provide an organization-wide involvement, people from all organizational levels should participate. For example, having the representatives of the top-level management, as well as people from the mid-level and operational level.

The technology transfer process itself is standardized by the GQM+Strategies training package\(^3\) and it involves an initialization step followed by a *characterization meeting* where the basic information about the method is provided to practitioners in order to better identify topics of interest and plan further steps. After that, the method *tutorial* is given to practitioners. The purpose of the tutorial is to transfer basic knowledge about the GQM+Strategies concepts and grid derivation process, followed by the exercise of developing a grid for their environment and their goals.

The grid development exercise requires several sessions and iterations for a successful completion. It usually starts with *preparation work*, which can be done off-line by a researcher or during a meeting with the practitioner who represents the method’s champion user. The method champions are practitioners, usually one champion per pilot, who are involved in all activities of the pilot from the beginning until the end. Champions receive additional training about the methodology because they are the main contact points and they also act as method promoters within organizations. After defining the starting points for developing a grid, *group sessions* are conducted for brainstorming goals and strategies. In between the sessions off-line work is done by the researcher, and if needed additional meetings and consultations with the method champion.

Finally, when the exercise is complete and the resulting grid is reviewed by all contributors, a joint *feedback session* is organized where the resulting grid is presented in order to invoke comments and discussions among practitioners from different organizational units. The purpose of the feedback session is to reach a common understanding about organizational goals and strategies.

Researchers can have several roles during the technology transfer process. They participate in preparing and planning the pilots. During the execution of pilots, they have an active role in knowledge transfer, while later, during the grid development exercises they play mainly the role of moderators, e.g. guiding discussions and helping practitioners to reach consensus about organizational goals and strategies.

\(^3\)http://www.iese.fraunhofer.de/en/products/gqm.html
4.1.2 **Data Collection**

The pre-study research goal of evaluating the GQM* Strategies approach in real world settings can be achieved by observing interactions between practitioners and the approach basically, studying their reactions while they are applying the approach in their environments and asking their opinions about the approach.

Such studies rely primarily on the feedback provided by people. Therefore, qualitative techniques were selected for data collection, in particular, participant observation (section 2.6.1) and interviewing (section 2.6.2).

**Participant Observation**

Throughout the study I was personally involved with the activity that was the object of the study. The activity observed was the GQM* Strategies grid development. During the activity, the interactions among participants allowed the observation of how people react to the grid development process and to the specific steps of the process. Besides the grid development activities, I also participated in different meetings with the industry representatives for preparing the GQM* Strategies pilots. Those meetings are also included in the field notes.

The pre-studies with the GQM* Strategies method had a general purpose to test the methodology in real world settings. Therefore, the field notes that were collected are also general in character. However, after several studies emerging patterns were discovered, and those observations are summarized and presented in this chapter (section 4.5).

The participant observation technique is very useful, but at the same time it can be very costly and effort intensive if robustness and high validity is required from the data collection process. Therefore, the participant observation technique was used in a combination with interviewing.

**Interviewing**

Interviewing was used in combination with observations, and early observations actually helped to define the focus points for the interviews. The interview guide that was designed is semi-structured and is given in Appendix 1. The questions were organized in five consecutive sections:
- **Warm-up questions**: demographic and context setting questions;
- **General questions**: questions addressing work responsibilities, typical problems, and the situation before the GQM+Strategies pilot;
- **Observation questions**: observations regarding the application of the method;
- **Follow-up questions**: more in detail elaboration of the method’s effects; and
- **Wrap-up questions**: closing questions.

The first two sections were about the organizational context and situation before the pilot. The observation questions probed interviewees’ opinions regarding whether the GQM+Strategies approach is beneficial. For example: *Was the process of building the grid helpful for your organization?* [The focus is on the process (activities, set of tools (e.g., templates), and additional materials).] This question was followed by the sub-questions: *What activities of building the grid did you find most useful? Why are those activities useful? What problems did you have with the grid building process?*

While the observation questions were used to identify the "topics" of interest, the follow-up questions tended to go into detail about those previously identified topics. For example: *How did the process of building the grid or the resulting grid help you to address the specific problems you mentioned?*

The wrap-up questions, section five, concluded the interview by asking interviewees to summarize what was and what was not good about the grid development. Also, in the last section, an opportunity was given to interviewees to discuss any issues that might have been missed during the interview.

The grids developed during pilot projects contain confidential and sensitive information that organizations may not be willing to make public. Interviews were an effective instrument for collecting and publishing data about grid development activities and the grids without revealing confidential information.

The guide was designed to keep the interview approximately one hour long. The assumption was that all interviews can be recorded and transcribed for later data analysis.

### 4.1.3 Data Analysis

The collected data were derived from two different cases. In those two cases, the technology transfer process was done by different people, which reduces the threat of having the same people doing the technology transfer and repeating all positive or negative actions from case to case. Therefore, the data analysis has elements of a
cross-case analysis [177]. It is not strict and formal as described in the literature, but it still provides useful comparisons between cases.

The codes and coding procedures are the main tools of the qualitative data analyses (section 2.6.4). The objectives of the pilots and the design of the interview questions led to an initial set of codes. The objective was to empirically evaluate the GQM+Strategies approach; therefore, understanding the effects of a grid development pilot was necessary. With the same purpose, the interview guide was designed to discuss the situation before the pilot and the situation after the pilot.

![Fig 20. The topics of interest organized in temporal sequence.](image)

The codes were organized in a temporal sequence by classifying the codes under one of three temporal codes. Figure 20 depicts the topics of interest organized in a temporal sequence, separating situations before, during, and after the GQM+Strategies pilot. The temporal sequence or timeline is coded with temporal codes that also act as the top-level codes, i.e. top-level nodes in a coding hierarchy. The temporal codes are:

1. [B4GQMS]: the characterization of organizations before grid development pilots,
2. [GRID DEV]: the actual process of developing grids, and
3. [EFFECTS]: The effects of the GQM+Strategies pilots on piloting organizations.

**Characterization** The codes for characterizing a situation in a case organization before the pilot are give in Table 9. The topics of interest are related to practices and activities that are connected with the problems that a successful organization needs to address (chapter 1). Those include questions about the organizational alignment, communication, measurement practices, and organizational reasoning.

Under the alignment, the specific actions that an organization is taking are considered in order to synchronize, i.e. align, different units of the organization. Such
synchronization requires specific knowledge of the organization or alignment knowledge. Alignment knowledge is used to develop organizational strategies and set goals for different units. Alignment knowledge needs to be communicated, shared, and made visible in order to have the desired impact on organizational activities and operations. Visibility or transparency is about the explicit communication of the organizational strategies or plans and making the alignment information accessible, i.e. increasing organizational transparency.

Furthermore, organizations need to have mechanisms to monitor and track whether goals and targets are achieved or not, and those mechanisms are based on or involve measurement. However, organizations need to reason about their goals and strategies, e.g., if those were the right goals and strategies, how much value they brought, and how much risk they contained.

Focusing on the use of other methods that are comparable with the GQM* Strategies approach is a useful source of information for learning through comparison with other methods. Exposing practical problems is important for understanding the practitioners’ motivation for applying the method in the first place. Interviewees were explicitly asked about the motivation for piloting the GQM* Strategies approach. The motivation also helps to expose the areas where the method is applicable.

**GQM* Strategies grid development pilot** The process of a grid development can be analyzed from several positions: adoption, applicability, and costs. The coding schema for a grid development activity is given in Table 10.

**Adoption** is about how well the approach is accepted by practitioners and the organization, i.e. to what extent the grid covers the organizational structure. The adoption process is initiated by knowledge transfer (technology transfer), but the level of adoption depends on existing skills and other approaches that are used by an organization. If the approach is perceived as useful by different groups of people inside the organization, it will have a higher chance of acceptance.

**Applicability** is really about classifying the type and size of the industry. The real costs of grid development and later maintenance are hard to identify at this early stage of method application. Therefore, the purpose of investigating costs is to understand the cost components, such as how much effort is put in by different people, how many group sessions are needed, etc.

33 addressing the same or similar problems as the GQM* Strategies
<table>
<thead>
<tr>
<th>Code</th>
<th>Sub-codes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALIGNMENT (AL)</td>
<td>DEV STRATEGY</td>
<td>Efforts to synchronize different organizational units to work toward same goals. Developing strategies or plans for achieving organizational goals and targets.</td>
</tr>
<tr>
<td>GOAL SETTING</td>
<td></td>
<td>Defining goals or targets for different organizational units</td>
</tr>
<tr>
<td>MEASUREMENT (M)</td>
<td>FOLLOW-UP</td>
<td>Follow up of organizational plans (i.e., strategies and goals)</td>
</tr>
<tr>
<td></td>
<td>DATA COLL</td>
<td>Collecting data for the purpose of monitoring and control operational activities</td>
</tr>
<tr>
<td>MOT</td>
<td></td>
<td>Motivation for using the GQM+ Strategies method</td>
</tr>
<tr>
<td>OTHER METHODS (OM)</td>
<td></td>
<td>Other methods used by organization for defining or measuring goals and strategies</td>
</tr>
<tr>
<td>PRACTICAL PROBLEMS (PP)</td>
<td></td>
<td>Problems and issues identified by practitioners</td>
</tr>
<tr>
<td>REASONING (RES)</td>
<td>ORG DEPENDENCIES</td>
<td>Analyzing dependencies between organizational units, between goals and strategies of different units</td>
</tr>
<tr>
<td></td>
<td>STRATEGY EFFECT</td>
<td>Analyzing the effectiveness of organizational strategies. Have they caused the desired effects? What kind of other effects they caused?</td>
</tr>
<tr>
<td></td>
<td>VALUE REALIZ</td>
<td>Value analysis and monitoring of organizational goals and strategies. What are the cost and benefits of strategies? Are strategies bringing perceived benefits?</td>
</tr>
<tr>
<td>VISIBILITY (VIS)</td>
<td>COMM</td>
<td>The practices for communicating and sharing organizational plans, goals, and strategies. Communicating organizational big picture.</td>
</tr>
<tr>
<td></td>
<td>TRANSPARENCY</td>
<td>The information accessibility regarding organizational big picture, i.e., plans, strategies, goals.</td>
</tr>
</tbody>
</table>
Table 10. The coding scheme for [GRID DEV].

<table>
<thead>
<tr>
<th>Code</th>
<th>Sub-codes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADOPT</td>
<td>GRIDCOV</td>
<td><strong>Grid coverage.</strong> To what extent organization was covered by a grid development pilot and which parts of organization should be involved in grid development activities.</td>
</tr>
<tr>
<td>INT</td>
<td></td>
<td>Integration of the GQM*Strategies approach with existing methods and practices within organization.</td>
</tr>
<tr>
<td>KTRANSFER</td>
<td></td>
<td>Knowledge transfer. How well did the practitioners understand the GQM*Strategies concepts? Are they capable of evolving and maintaining the resulted grid?</td>
</tr>
<tr>
<td>USERS</td>
<td></td>
<td>The users of the grid. Practitioners who were involved with a grid development but also other users who might benefit from the grid.</td>
</tr>
<tr>
<td>APP</td>
<td></td>
<td>The method’s application domain, e.g., industry type and size. Additionally, for which problem areas the method was used.</td>
</tr>
<tr>
<td>COST</td>
<td></td>
<td>The cost of a grid development activity. Effort descriptions and estimates, resources used, etc.</td>
</tr>
</tbody>
</table>

**Effects**  
The motivation for testing the GQM*Strategies approach is the practitioners’ belief that the approach can help them in solving their problems. The codes for effects are given in Table 11. After the pilot, the practitioners are in a position to discuss the effects of the pilot, whether the approach succeeded in addressing their problems, and if not, why not. The codes for effects are given in Table 11.

The first one represents the **impact** of a developed grid on issues and problems that were identified before the GQM*Strategies pilot. These mainly correspond to the topics used for the characterization of the situation. In contrast, the second one represents **obstacles** or challenges that emerged during pilots.

The effects include the suggestions on how to improve the approach or the GQM*Strategies grid development process, as well as practitioners’ wishes or “nice to see” future features. They all represent **opportunities** for further development of the approach.
<table>
<thead>
<tr>
<th>Code</th>
<th>Sub-codes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMPACT (IMP)</td>
<td>ADVANT</td>
<td>Advantages of the GQM+Strategies over other similar approaches used by an organization</td>
</tr>
<tr>
<td></td>
<td>ALIGN</td>
<td>Ability of the grid development process to align different organizational units and connect and synchronize organizational goals and strategies from different organizational units.</td>
</tr>
<tr>
<td></td>
<td>COMM</td>
<td>Use of the grid and grid development process as an organizational tool for communicating and sharing plans and strategies across different organizational levels.</td>
</tr>
<tr>
<td></td>
<td>MEASURE</td>
<td>Impact of grid developing activities on organizational measurement practices.</td>
</tr>
<tr>
<td></td>
<td>MONITOR</td>
<td>Impact of grid on monitoring the implementation of organizational goals and strategies.</td>
</tr>
<tr>
<td></td>
<td>ORG REAS</td>
<td>Use of a grid for organizational reasoning about goals and strategies. (see code [REASONING])</td>
</tr>
<tr>
<td></td>
<td>TRANSP</td>
<td>Ability of the grid definition phase to improve organizational transparency and make visible the organizational interdependencies.</td>
</tr>
<tr>
<td></td>
<td>USEFUL</td>
<td>Usefulness of developed grid and GQM+Strategies approach for an organization.</td>
</tr>
<tr>
<td>OBSTACLES (OBST)</td>
<td>OCOMM</td>
<td>Challenges in communicating GQM+Strategies concepts or presenting grids, including language issues, e.g. business people use different language than developers</td>
</tr>
<tr>
<td></td>
<td>OCOMP</td>
<td>Problems with dealing with complex grid structures or an expressed concern regarding a potential grid complexity.</td>
</tr>
<tr>
<td></td>
<td>OTOOL</td>
<td>Problems with the GQM+Strategies visualization tool</td>
</tr>
<tr>
<td></td>
<td>VISUAL</td>
<td>Challenges and requirements for a graphical representation of the GQM+Strategies grid</td>
</tr>
<tr>
<td></td>
<td>MAINT</td>
<td>Challenges or expressed concerns regarding the maintainability of the GQM+Strategies grid</td>
</tr>
<tr>
<td></td>
<td>LEARNING</td>
<td>Impact of the learning curve on full utilization of the GQM+Strategies approach</td>
</tr>
<tr>
<td>OPPORT (OPP)</td>
<td>SUGG</td>
<td>Suggestions for improving the GQM+Strategies approach and grid development activities</td>
</tr>
<tr>
<td></td>
<td>WLIKE</td>
<td>A wish list. What would practitioners like to have or see as a part of the GQM+Strategies methodology?</td>
</tr>
</tbody>
</table>
4.2 GQM$^+$Strategies Industry Pilots

During the time period from 2008 to 2011, the opportunity arose to observe or interview and gain feedback from several industrial audiences. These opportunities took several forms. There were high-level presentations about the GQM$^+$Strategies method that involved wide audiences from the industry. Such presentations had the purpose of communicating the basic ideas of the method. Tutorials were given as training in the method tutorial (usually one-day long training). Only organizations that were committed to complete a grid development exercise, are considered as industry pilots.

Figure 21 shows the overview of industry pilots and presentations in which I personally had some involvement. In 2008, a first industry pilot was done with a company in the telecommunication network testing business (Telco#1). The company is developing software solutions for testing telecom networks. A year later, a pilot was run with another company that is developing embedded systems for the telecommunication networks (EmbSys#1). At the end of 2010 and the beginning of 2011, I was involved with developing a product strategy using the GQM$^+$Strategies approach with another company (Telco#2) that has a similar business as the first organization, i.e. network testing solutions. In the fourth case (ITDep#1) the grid piloting was done by other researchers, and practitioners were interviewed about the effects of the method application.

Besides the industry pilots, several method presentations (Soft#1 and Soft#2) were given to different audiences, including management and quality assurance and metrics people. Although the grid development exercise were not conducted, the reactions and feedback regarding the GQM$^+$Strategies method were constructive and useful.

For collecting feedback from pilots, several data collection mechanism were used, either individually or in combination. Those mechanisms were: observation, interviews, and surveys. The data collection mechanisms are marked in the last three rows in Figure 21.

For investigating the practical problems and evaluating GQM$^+$Strategies pilots, the interview was the data collection mechanism of choice. Therefore, the following section will discuss pilots that used interviews as a data collection mechanism. In addition, the interview findings correspond well to the feedback obtained from other practitioners during the method pilots.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Method presentation</td>
<td>◆</td>
<td>◆</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Method tutorial</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Grid development exercise</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Grid pilot</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Participant observation</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Interview</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
<tr>
<td>Survey/Questionnaire</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
<td>★</td>
</tr>
</tbody>
</table>

Legend:
- ★ Activities where I participated
- ◆ Activities conducted by others

Fig 21. Overview of the industry involvement with the GQM+ Strategies activities.

4.2.1 Industry Case 1: Telco#2

The grid piloting was done within a division of a large international ICT company with several branches worldwide. The company has been present in the telecommunications market for almost two decades. The primary products of the company are telecommunication network tools and equipment. The organizational unit where this study was carried out is located in Finland, and the entire site employs over 100 people. Overall, 10 people participated in the study, involving representatives from different management and operational levels. And, two organizational units, the sales and R&D departments, were involved in developing the GQM+ Strategies grid.

The objective of the pilot was to implement the GQM+ Strategies with the purpose of developing product strategy and linking the product strategy with top-level corporate goals.

Before the grid development workshop, a two-hour tutorial was given to participants. After the tutorial, the first workshop was used to clarify top-level organizational goals and strategies (goals defined by corporate headquarters). What followed that was a brainstorming session with representatives of the sales and R&D teams.
From the beginning of the pilot, the main sponsor for the work was at the same time the method’s champion user, and he was the representative of the top-level management for that corporate site.

After the first workshop, another brainstorming session was held by the product development team (R&D), which resulted in a clear specification of the constraints, key context and assumption elements, and some draft goals and strategies. Those goals and strategies were further refined and formalized with a GQM+Strategies template during several individual meetings with the method’s champion user. Each of those meetings was followed by several hours of off-line work.

The company had defined strategic goals and objectives, which were the starting point to derive the grid. The top-level business goal was defined as: activity to increase (focus) the sales for (object) all product lines and all regions for the magnitude of 20%\textsuperscript{34} CAGR; within a timeframe of three years in the context (scope) of all corporate site with constraints of current resources availability. Relations to the corporate policy.

One of the strategies employed to address the upper goal was to utilize a new coming technology for increasing the competitive advantage of current products.

The resulting grid contained detailed goals and strategies for the product development team (i.e. product strategy), with documented context and assumptions for goals and strategies, and how those goals link to other strategies (sales strategy and organizational top-level strategy).

The method’s champion was a representative of the top-level management, and his personal interest in the GQM+Strategies approach was motivated by managerial issues. Therefore, his perspective and feedback on the pilot was from the management point of view.

4.2.2 Industry Case 2: ITDep#1

For the second pilot, I was not personally involved with the technology transfer or grid development activities; therefore, details about those activities are not provided.

The second pilot was done within an IT department of an international company. The company is one of the largest mining companies in the world and the largest company in South America. The IT department employs roughly 140 people organized in several groups addressing different IT-related tasks, such as classical IT services, services

\textsuperscript{34}Due to confidentiality reasons numbers are not real.
related to certain information units that have to be maintained, software and system
security, system architecture, and a Software Factory that is responsible for maintaining
and integrating existing systems as well as developing new IT-based software systems.

The top-level goal of the ITDep#1 was defined as: activity to improve and maintain
(focus) the information quality for (object) information units of all business processes for
the magnitude of 10% decrease of #reds and 10% decrease of #yellows for each attribute;
within a timeframe every three months in the context (scope) of information division for
the upstream Start with exploration unit data; Start with attributes uniqueness and
consistency with constraints of use of resources for other activities.

In order to address the IT department top-level goal four strategies were defined. The
grid piloting was done in 2009 and 2010. After, the successful technology transfer, the
local resources (method champions) were created within the ITDep#1, who continued
to work on the grid development and strategies implementation. In 2011, those two
method champions were interviewed about their experiences with the GQM* Strategies
approach.

4.3 Data Collection

For the first pilot only the method champion was interviewed, using the interview guide
(Appendix 1). The interview was conducted at the researcher’s site, in order to be
able to record the interview (the use of recording devices at organizational premises
requires a special permission). The language of the interview was English. The interview
recordings were transcribed.

For the second pilot, two interviews were conducted over Skype. In the first interview,
two interviewers were present, one having a leading role and the second taking notes.
For the second interview, a third researcher participated as a note taker. These interviews
were also recorded and transcribed.

4.4 Data Analysis and Findings

The interview transcripts were coded using the coding scheme (defined in section
4.1.3), and here the most interesting findings were selected from the research point of
view. The data analysis process was very close to data; in other words, there were no
theories generated from the data. At this stage of the research process, i.e. discovering
the practical problems and motivating research questions, that kind of analysis was sufficient.

For analyzing interview transcripts the NVivo software package was used, along with its support for open coding activities and later for analyzing references during axial and selective coding.

What follows will discuss the most interesting findings.

4.4.1 Finding #1: disjoint goals and strategies are posing a real challenge for the organizational alignment and transparency.

Planning and developing organizational strategies requires specific cognitive activities and strategic thinking. Such strategic thinking is contained at the upper levels of an organization and is disconnected from the rest of the organization. The effects of the existing disconnect, i.e. misalignment, are (1) irresolute actions that mismatch strategic goals, and (2) inadequate organizational guidelines that could be followed at the operational level. As noticed by an interviewee who has a high management position in his organization:

Another thing in the current practices in the strategic level is that there seems to be no, there is no direct, let's say, product strategies. It's division-level strategies and that's it. So there's kind of a link missing in the strategy. There's a strategy in the high level, strategy in the division level and then there are the product, program, those reviews. So it's, there's no really pure strategy put down to the product level what we can then follow, let's say the longer-term strategy. [B4GQMS, PP, AL, M, VIS]

Furthermore, the organizational transparency is important for getting everybody into the same strategic direction; however, the representatives of management expressed serious concerns of not being able to follow through to strategies at operational-level activities and to see how those activities contribute to the organizational goals. At the same time, the people from lower levels have difficulties in reaching an understanding and aligning with the higher organizational goals and strategies, i.e.

35Name, and version, web link
I didn’t have any tools for reaching our goals or the goals of project to the goals of the company. And actually the same situation was for any other project inside the IT department. [B4GQMS, PP, MOT, VIS]

This seemingly simple issue of linking and connecting activities from different levels was accentuated by all interviewees as a real challenge for their organizations.

4.4.2 Finding #2: the GQM\textsuperscript{\textregistered} Strategies grid development process helps identify and resolve the issue of organizational misalignment.

That existing gap between higher-level organizational strategies and lower-level organizational goals was the main motivator for piloting the GQM\textsuperscript{\textregistered}Strategies approach.

The GQM\textsuperscript{\textregistered}Strategies grid development pilots demonstrated the method’s capability to further develop higher-level organizational strategies by involving people and stakeholders from lower levels and incorporating their understanding of the best way to address the organizational goals and strategies, as well as communicating feedback to the upper-level management. The feedback provides a better insight regarding the effects of the upper-level strategies on operational activities:

I found the approach very good, because then you have to stop thinking about the strategy and the goal of this [level], and what is the meaning [(i.e. effects)] at the bottom level. [EFFECTS, IMP, ALIGN, USEFUL]

The whole GQM\textsuperscript{\textregistered}Strategies process helped to align views of the people from different organizational units through joint workshops and discussions:

I think the, for example, even the brainstorming itself was useful, and the kind of work together with the different units, like sales and there was marketing and there was R&D and there was PLM. So that work, even though there was not that much of them [participants], this was useful definitely, to find out and kind of clarify all of us the same view where we are going to. This was very useful. And then, kind of the ideology or our finding out the path
to my main interest, to the product strategy, from the corporate strategy, this was very good for me to understand the bigger picture. [EFFECTS, IMP, ALIGN, USEFUL, TRANSP, COMM]

Furthermore, those group sessions helped to identify gaps or issues of misalignment, and afterwards the GQM+ Strategies approach provided tools to address those gaps, i.e. define the missing goals and strategies.

4.4.3 Finding #3: organizational goals whose values are not expressed by financial indicators tend to become less important than the goals whose values are quantified by financial indicators.

The missing link between different organizational levels makes the value contributions from different projects inside an organization unclear. Different groups inside an organization define their own goals with the implicit belief that their goals are valuable for the business. However, they do not understand how the goals of other units contribute to the business goals, how their goals affect other projects, and how other projects may contribute to them. For example, the practitioners expressed the need for having a mechanism to make value contributions visible:

So, we want to make explicit our contribution to the business, because there was really concern from the CEO, he used to question the architectural team, because he said architecture is too abstract, and they don’t really, I don’t really see a contribution. [B4GQM, PP]

The lack of such a visibility can cause uninformed decisions at higher levels of the organizations.

Furthermore, it is a very common practice to follow-up the financial indicators at higher organizational levels, and those financial targets become goals for themselves, where the other organizational goals are slowly forgotten due to the lack of mechanisms to follow them in the same way as financial indicators.
At strategic level we define targets, numbers. It’s purely a euro-based follow-up. Not really following the strategic goals. [B4GQMS, M, FOLLOW-UP, PP]

All interviewees were in agreement that additional information about the (estimated) costs and benefits of the goals can significantly increase the visibility of the value contributions from different organizational units:

Do you think that information about the cost of the strategy, estimated cost and estimated benefits, would help when higher-level decisions are made?
Yeah, of course. It's very important to know the cost of it. If I want to enhance one indicator, how much will it cost, that's very important. [EFFECTS, OPP]

Although, the GQM*Strategies approach can make the links between different organizational goals visible there is a further opportunity to incorporate cost and benefit estimates of the goals and strategies, i.e. in a way to quantify the value contributions of different organizational units.

4.4.4 Finding #4: the assumption elements are useful in coping with the complexity and uncertainty of the business environment.

The external and internal environment of ICT companies is complex and dynamic. The dynamism of the environment brings additional uncertainty to the strategies development process.

The practitioners found the assumption elements of the GQM*Strategies grid particularly useful for handling the uncertainty and complexity of the environment:

I like very much for example the part of having the assumptions there, to clarify this, that we are explaining outcome with assumptions. So this was a very good way to handle the complexity of it. [EFFECTS, USEFUL]

The use of context and assumption elements significantly increases the understanding of the background rationales for strategies and goals.
4.4.5 **Finding #5: better mechanisms for incorporating the risk information into the GQM*Strategies grids are needed.**

For stakeholders to fully anticipate risks they need to understand the environment with its existing complexity of different events that might occur. However, in large organizations analyzing the environment to that level of understanding requires the involvement of a whole team of knowledgeable representatives from different organizational levels. Therefore, the risks need to be characterized (e.g. risk exposure, impact) in order to be properly communicated to all stakeholders.

They [people involved in strategy development] should see the main risks that are compromising their strategies and goals. And this they cannot see in the, here in the graph [the GQM+Strategies grid]. So maybe that could be interesting to see in the future. [EFFECTS, SUGG, WLKE]

Although, the risk can be addressed through the GQM*Strategies assumptions, it can be improved with additional mechanisms to analyze and quantify the risks of goals and strategies.

4.4.6 **Finding #6: a mechanism for quantifying the effectiveness of strategies is needed.**

In situations when several strategies are considered and/or used for a set of goals, different criteria can be defined to analyze the value of strategies. One such criteria is to analyze the effects of the strategies on goals. Consequently, strategies with larger effects (i.e. more effective strategies) would be considered more valuable or important for some goals in a grid.

I think the GQM+Strategies does not have a measure for saying how much does one strategy impact the goal, does it? How a strategy impacts one goal, I would say, okay, this is the strategy, it impacts 30 percent the goal and 30 percent some other goals. [EFFECTS, OPP, SUGG]

Such assessments could be provided by experts during the grid development:
[when defining a strategy for a] goal [an] expert decision [is made] on how much would this strategy contribute to the goal. [EFFECTS, OPP, SUGG]

The interviewees from both organizations suggested that a mechanism for quantifying strategies’ effects on goals would make the GQM*Strategies approach more attractive.

4.5 Reflections and Observations

Four GQM*Strategies pilots were conducted from 2008 to 2011. All those pilots involved technology transfer, such as tutorials and workshops, followed by an exercise in grid development. The overview of all activities that were done within the industry is given in Figure 21. The data collection and analysis methodology varied depending on the pilots’ objectives and research goals.

The following paragraphs will summarize the observations regarding the GQM*Strategies technology transfer and grid development activities.

Observation #1: the GQM*Strategies grid development was accompanied by discussions about the value of goals and strategies. The practitioners engaged in discussions, i.e. informal analysis, about the costs and benefits of goals and strategies either while the grid was developed or afterward, during a grid feedback session. They suggested that after investing the effort to systematically document the goals and strategies, an additional effort to analyze cost and benefits of those goals and strategies would be acceptable. The extra effort would pay off several times and the whole resulting structure would be more attractive for business managers.

Observation #2: the current GQM*Strategies mechanisms are not sufficient for the effective handling of multiple strategies of a single goal. Often during the workshops, several good strategies for a single goal emerged. Such situations were resolved by determining which strategy was the most promising (the most effective under the given circumstances). However, even more complicated cases can occur, e.g. when several strategies together can be more effective than an individual most effective strategy.

Observation #3: the adoption of the GQM*Strategies approach largely depends on the pre-existing organizational skills and practices. In all pilots,
the practitioners had a positive attitude toward the GQM+Strategies approach, meaning that they understood the potential benefits that the approach could bring them; otherwise, they would not have allowed pilots. All pilots required effort for explaining the potential benefits; however, there were pilots for which all the benefits that the approach could provide could not be demonstrated. The reasons for this discrepancy were different starting points with respect to certain organizational skills. For example, if an organization is using goal-driven measurement approaches, then the implementation of the GQM+Strategies is much easier than if the organization does not have any good measurement practices in place.

Observation #4: the contextual information shapes the entire GQM+Strategies grid and extracting the relevant contextual information is important for the quality of grids. The high-quality grids capture the relevant information for a particular organization. However, that information is only known to people from that organization. When the process of extracting information is done as a group exercise (usually done as a group brainstorming session), the moderator of the exercise needs to be highly concentrated and focused on determining what information is relevant. Maintaining that level of attention and focus for several hours is demanding.

Observation #5: the grids successfully conveyed operation-level issues and the effects of the upper-level goals and strategies back to the upper-level management and vice versa. In pilots where representatives from all levels were involved, the top-management was surprised when they realized the effects of their goals and strategies throughout the organization. The GQM+Strategies grid helped them to better understand the specific organizational issues.

4.6 Research Questions

The feedback from the GQM+Strategies pilots was used to identify research problems and, thus, research opportunities. Many opportunities for method improvement were suggested by the observations and interviews. However, a subset of these were chosen that are related to the research problem areas identified in Chapter 1, i.e. value-based software engineering (section 1.1.1) and goal-driven alignment (section 1.1.2). The collection of problems that were selected forms a logically connected and interdependent package of research opportunities as articulated below. These research questions will
act as a set of requirements for the research. This selection should have industrial impact while also representing an intellectually interesting set of research problems. Furthermore, the opportunities that have not been addressed in this dissertation will be discussed in the future work section.

4.6.1 **RQ1: How can the GQM* Strategies approach be used to document value propositions across an organization?**

The seemingly simple problem of connecting or linking the value propositions (i.e. goals and strategies) of stakeholders at different organizational levels represents a real challenge (Finding #1; Observation #1) for the creative industry today. The GQM* Strategies can help link and connect goals from different organizational levels (Finding #2). Therefore, it is an interesting research question to explore further how the GQM* Strategies can provide a notation for enunciating an organizational value proposition.

4.6.2 **RQ2: How can the GQM* Strategies be instrumented to monitor and control a value creation process?**

The GQM* Strategies grid structure is suitable for performing cost/benefit analysis (Observation #1). Having such capabilities within the GQM* Strategies approach was perceived as an important extension; otherwise, important goals that are not financially tracked might fade out over time (Finding #3). This issue poses a research sub-question: **RQ2.1: How can a cost/benefit analysis of the GQM* Strategies grid be performed?**

The GQM* Strategies approach allows setting of the assumptions about the real world or business environment that helps in communicating uncertainties that impact goal definition and strategy selection (Finding #4). Furthermore, it seems to be a very common situation in software-intensive organizations that stakeholders at one level do not have a full understanding of issues from other levels (Observation #5), and because of that they are not capable of understanding risks concerning goals and strategies other then their own. Therefore, it is important that risk information is communicated together with goals and strategies to all stakeholders (Finding #5). That is a motivation for the research sub-question:
**RQ2.2:** *How can risks regarding value materialization be identified?*

**4.6.3 RQ3:** *How can the GQM*\(^*\) Strategies be used to analyze dependences between different value propositions across an organization?*

Different strategies can have different effects on the same value proposition (Observation #2), e.g. two less promising strategies can be more effective when applied together than one single strategy. Therefore, there are different ways of evaluating the contributions of a particular strategy. Some value propositions resulting from one strategy can have a negative effect on other value propositions, which is usually manifested as contradicting goals. Therefore, it is important to provide a mechanism for analyzing and quantifying goal–strategy dependences (Finding #6; Observations #5, #2):

**RQ3.1:** *How can the goal–strategy dependences in the GQM*\(^*\) Strategies grid be quantified?*

Furthermore, such a mechanism would be particularly useful for the further analysis of the risk impact on other goals and strategies in the GQM*\(^*\) Strategies grid, which leads to the following research sub-question:

**RQ3.2:** *How can the impact of the risky value propositions on other value propositions within an organization be assessed?*
Following the research approach described in Chapter 3, the research process goes through two phases. In the first phase, practical problems and improvement opportunities are identified during the pre-study empirical investigations of the GQM*Strategies approach (Chapter 4). Those practical problems motivate research questions and serve as requirements for the future solution (section 3.2.1). The second phase is the solution development process that consists of activities that are associated to design hypotheses and the research problem (section 3.2.2). The purpose of the research problem and design hypotheses formulation is to ground the empirically motivated practical problems within specific research areas in theory. The research problem represents a topic that identifies and bounds the body of knowledge to which the contributions, current and future, are made. The concrete contributions represent solutions developed to evaluate design hypotheses. In the following section the research problem and design hypotheses that clearly link the practical problems and gaps in existing theory will be formulated.

5.1 Research Problem

The research questions (section 4.6) are about: enunciating organizational value propositions (RQ1), monitoring organizational value creation (RQ2), and analyzing and quantifying dependences between value propositions (RQ3). However, the research questions relate to the research problem areas (Chapter 1) and the top-level concerns that a successful organization needs to address (ACMR cycle in Chapter 1).

Supporting organizations with a mechanism to document value propositions of different organizational units (RQ1), also helps them to better understand how organizational units can be aligned in a way to maximize overall value creation. Making organizational value propositions explicit and documented together with all relationships among different organizational levels and units helps in communicating the organizational big picture and increases the transparency within the organizational.
The organizational value management practices rely on and use value monitoring mechanisms that deploy various measurement schemes in order to collect and interpret data about the value creation. Furthermore, the deployment of such measurement programs brings additional clarity to the organizational activities and operations and directs the attention of people to value critical issues and activities. Therefore, the second research question (RQ2) addresses organizational measurement, and to some extent, organizational communication problems.

The organizational reasoning about goals and strategies depends on the understanding of dependences among different organizational units. The quantification of the value proposition dependences of those organizational units (RQ3) can significantly increase the quality of the reasoning about organizational goals and strategies. However, people need to share that understanding and knowledge. One mechanism that can help with sharing that knowledge is quantification of the value propositions. Also, organizational reasoning includes a form of reflective thinking about the effectiveness of strategies in addressing different goals, and that kind of thinking helps to better understand the alignment and how the organizational units can be aligned based on past experience and performance.

Some mechanisms address parts of these problems, e.g., value-based software engineering (section 2.4), or GQM+Strategies (section 2.3). However, they fall short of implementing a complete cycle of goal alignment and sharing (communication), measurement, and reasoning through the evaluation and improvement of organizational goals and strategies.

Value-based software engineering has mechanisms to analyze value propositions and, to some extent (at the project level, but not entirely at an organizational level), to control and monitor value realization (see section 1.1.1). It does not address together the organizational alignment and measurement issues. The GQM+Strategies approach was designed to address issues of organizational alignment and measurement together; but, it lacks mechanisms for value and risk analysis. Both approaches lack capabilities to support organizational reasoning about goals and strategies, or to address how the success or failure of different organizational units and projects within those units affect the whole organization, i.e. other units and projects within the organization.

The aim of this work is to further evolve the GQM+Strategies methodology in order to answer the research questions by providing a comprehensive solution and to generate a cycle of activities that will allow an organization to evaluate and assess the values and risks of its aligned strategies and goals. Therefore, the research problem can be
formulated as follows:

Research problem: Assuming that an organization successfully uses the GQM+Strategies approach, extend the approach with other forms of analysis and reasoning that will help organizations better align and manage value propositions at different organizational levels.

In what follows, a set of design hypotheses and research activities are defined.

### 5.2 Design Hypotheses and Activities

The design hypotheses are theoretical constructs that propose which theoretical concepts or theories may be used, together or individually, to produce desired outcomes or results. Each hypothesis addresses one or more research questions, and those relationships are explained in the following sections. Developing a solution requires the selection of certain alternatives from a solution space (section 3.2.2). Using design hypotheses as theoretical guidelines helps to limit the potential alternatives. The development of the solution is a process, and such a process is conveniently described as a set of activities.

![Fig 22. The schematic view of the solution development process.](image-url)
Figure 22 gives a simplified schematic view of the solution development process. The simplification is due to excluding the time perspective and the activities’ iterations. The research was done over a longer period of time, and research activities were conducted in different time moments. The results of one activity impacted or were used by another activity. Nevertheless, the schematic overview illustrates the connections among research questions, design hypotheses, and activities.

5.2.1 Design Hypotheses

$DH_1$: Applying the GQM*Strategies notation to VBSE provides a means to enunciate stakeholders’ value propositions in a precise and testable format at different organizational levels in the form of goals and strategies.

Design Rationales The value-based software engineering has a well established empirical research community with an excellent publishing record (section 2.4). Furthermore, the well defined VBSE framework specifies criteria (requirements) to be met by an engineering method that aims at dealing with the value related concerns. Therefore, the VBSE framework is a good candidate for a reference model. In addition, the VBSE deals with multiple stakeholders at various organizational levels, and in similar way the GQM*Strategies integrates different stakeholders from an organization.

The design hypothesis $DH_1$ explores how the GQM*Strategies concepts, the meta-model and the grid derivation process (section 2.3), can be used to perform the VBSE activities (section 2.4). The result of merging the concepts of the VBSE and the GQM*Strategies is a goal-driven notation for documenting stakeholders’ value propositions across an organization. This is achieved by using the VBSE framework and embedding its key activities into the GQM*Strategies grid derivation process. The resulting process is capable of developing measurement schemes that are aligned with the organizational goals and strategies, i.e. the goals and strategies that stakeholders believe to be valuable for the business.

The hypothesis directly addresses research question RQ1 (Figure 22). Here the notation is based on GQM*Strategies concepts, such as goals and strategies, for constructing a single structure (i.e. the GQM*Strategies grid) that is capable of documenting value propositions of different stakeholders across an organization in a systematic way. Also, the grid development process can provide support for conducting
the key activities of the value-based software engineering. This hypothesis is also important for research question RQ2 in the following way. Having a notation whose end result is a single structure that documents and captures organizational value propositions while performing the VBSE activities that involve cost, benefits, and risk analysis of the GQM* Strategies elements is a good starting point for setting baselines for the value monitoring and control practices.

**DH2:** Applying earned value management (VBSE) concepts to the GQM*Strategies grid elements enables monitoring of organizational earned value through multiple stakeholder perspectives that allows effective cost management and the realization of benefits during the implementation of organizational goals and strategies.

**Design Rationales** Applying earned value management concepts is necessary in order to fully comply with the VBSE framework. A unique GQM*Strategies grid structure allows for connecting and integrating different organizational projects that use earned value management for monitoring and controlling implementation of goals and strategies of those organizational projects.

The design hypothesis DH2 investigates how the earned value management indicators (section 2.4.3) can be defined for monitoring the implementation of the organizational goals and strategies. The goals and strategies are specified and formalized with the GQM*Strategies grid. The earned value management provides several mechanisms that allow projects to monitor and control their schedule and cost with a set of earned value indicators. However, those mechanisms are contained within the project’s lifespan. Implementation of organizational strategies is more complex. The complexity is that multiple projects are running at different organizational levels within various organizational units, all those projects have different lifespans, and they are all connected and interdependent. The GQM* Strategies approach provides a means to deal with that complexity, and including earned value management practices provides not just a set of earned value indicators for monitoring cost and schedule, but also indicators for monitoring the realization of benefits across an organization.

This hypothesis is addressing research question RQ2 (Figure 22). The research question is concerned with the dynamics of the value creation process and earned value indicators are designed with exactly that purpose.
Applying causal reasoning (causality theory) to GQM*Strategies allows for a construction of quantitative goal–strategies–goals models that are useful for analyzing the impact of risky goals on other goals in a grid.

Design Rationales  The GQM*Strategies models are subjective, meaning that goal owners (people participating in the grid development) define goals and shape strategies according to their knowledge and understanding of the real world and organizational context. The most prominent reasoning framework that incorporates experts opinions and belief is causal reasoning or Bayesian belief reasoning [185]. Using causality theory provides the necessary formalism for the construction of quantitative goal–strategies–goals models.

The design hypothesis $DH_3$ examines what the GQM*Strategies grid structure represents as a whole for the organization. The grid is comprised of different organizational goals and strategies, and their links. In other words, the GQM*Strategies grid enunciates organizational or business goals and ways of achieving them as articulated by strategies at different organizational levels. For example, top-level goals define a future situation that should be the effect of the successful execution of the top-level organizational strategies. Lower-level goals are defined with the intention of having a causal effect on the upper level-goals, and so on. The whole process of deriving the GQM*Strategies grid embeds causal thinking. Therefore, causality theory can be used for analyzing dependences and relations between goals and strategies.

This hypothesis addresses research questions RQ1, RQ2, and RQ3. Research question RQ1 is concerned with documenting "individual" value propositions, but it is also concerned with compiling an organizational value proposition that includes those individual value propositions. Understanding what the GQM*Strategies grid represents as a whole helps compile such an organizational value proposition. Research questions RQ1 and RQ2 address the cost, benefits, and risk of the GQM*Strategies goals and strategies from two perspectives. The first is how to analyze each of the grid components (RQ1), and the second is how to monitor the progress and changes in estimated cost and benefits of each grid component (RQ2). The ability to quantify dependences between the grid components, i.e. goals and strategies, directly answers research question RQ3 and leads to better and more realistic control mechanisms of the cost, benefits, and risk.

The evaluation of the hypotheses leads to the following activities.
5.2.2 Activity 1: adding a value-based component to the GQM+Strategies approach

Description. Add a value-based component to the GQM+Strategies that provides a means of defining (enunciating) stakeholders’ value propositions at different levels in the form of goals and strategies. The expected result of the activity is a goal-driven VBSE notation. This activity is associated with design hypothesis $DH_1$.

Evaluation. As a proof of existence, a solution that evolves the GQM+Strategies to integrate the key elements of the VBSE framework is developed. The solution evaluation is analytical and involves the conceptual analysis of the VBSE framework and proposed solution.

5.2.3 Activity 2: building a model for monitoring cost, benefits, and risk of the grid elements

Description. Build a model for monitoring cost, benefits, and risk of the GQM+Strategies goals and strategies. The expected result of the activity is a GQM+Strategies goals and strategies value and risk-monitoring method. This activity is associated with design hypothesis $DH_2$.

Evaluation. As a proof of concept, a solution, i.e. a model and a set of earned value metrics for monitoring an organizational earned value is developed. The solution is evaluated through an example.

5.2.4 Activity 3: constructing the goals–strategies dependences models of the GQM+Strategies grids

Description. Develop a mechanism to analyze dependences between goals and strategies in a GQM+Strategies grid and quantify the impacts of risky goals on other goals in the grid. The expected result of the activity is a method for analyzing GQM+Strategies grid embedded risk and goal–strategy dependencies. This activity is associated with design hypothesis $DH_3$. 
**Evaluation.** A formal mathematical notation of the GQM*Strategies* concepts is developed as the basis for causal reasoning. As an existence proof, a goal–strategies dependency model is developed for the GQM*Strategies* grid. The solution is evaluated by piloting it and assessing the feasibility of constructing such goal–strategies dependency models. During the pilots, practitioners are exposed to a familiar GQM*Strategies* grid and asked to provide the necessary inputs for constructing goal–strategies dependency models. The models are used to analyze the size of a risk embedded in the grids. The results of the analysis are presented to the practitioners in order to gain an expert judgment on how well the results reflect real world situations.
III Developing the Solution
6 Organizational Value Alignment and Measurement with the GQM+Strategies

You can’t control what you can’t measure.
by Tom DeMarco
(Controlling Software Projects: Management, Measurement, and Estimates; 1986)

This chapter describes how the GQM+Strategies approach can be used to support organizational value alignment. Such support requires the integration of business value analysis with the GQM+Strategies grid derivation process.

Business value analysis attempts to explore the factors that shape the future instead of attempting to forecast the future. Business value analysis, identifies one clearly defined value from all stakeholders’ perspectives and quantifies the value of different options. Furthermore, a good value analysis provides metrics to capture value, control risks, and capitalize on opportunities during projects (Section 1.1). However, aligning organizational units to create value at the enterprise level gets less attention, in theory and practice, then creating value at the business unit level [120, p. 3]. Kaplan & Norton [120] observed that for a corporation to add value to its collection of business units and services it needs to align those units and create synergy, i.e. value added at the corporate level.

One of the results of the GQM+Strategies pilots (Chapter 4) was that understanding the value and contributions of the individual goals and strategies is important for several reasons. For example, one finding was that goals that are not quantified in terms of financial indicators, with respect to their cost and benefits, tend to fade over time, and people start to pay less attention to such goals.

In the software industry, where intangible products are produced by intangible production lines "operated" by creative people, the value alignment of different organizational activities is even more challenging than in the manufacturing industry. In order to integrate and align a variety of the activities of software product development under the perspective of value, a wider framework is necessary. Currently, the most prominent framework that addresses the value perspective of software development is value-based software engineering (Section 2.4).

The VBSE framework integrates all aspects of the software creation process into the perspective of value. Business or organizational value cannot be measured in the
same way as time or volume, because value is always context-dependent [27]. Added organizational value is not explicitly addressed in software engineering standards or its body of knowledge. The VBSE approach helps elicit sources of value, reconciling value conflicts and organizing activities in software engineering according to their value [30].

Boehm & Huang [43] emphasized that integrating value aspects in all phases of software engineering is important, but recognize that integrating the financial instruments in order to enable the value-based monitoring of processes is a challenge.

There are other approaches in software engineering that emphasize value, especially customer value, like agile approaches [61, 76] or lean software development [130, 150]. All agile approaches try to organize the product creation process in relatively short iterations where each iteration is supposed to deliver functionalities that are believed to have high customer value. The lean approach was imported from the manufacturing industry. One of the fundamental ideas of lean production is the optimization of the flow of value through the entire manufacturing, i.e. software development, process. However, the understanding of what lean software development is, is still vague. One possible perception of lean software development is that it is a vessel for scaling agile development from a project-level approach to an organization-wide approach [132].

Value-based software engineering has characteristics of a more general framework. This framework defines criteria and argumentation for those criteria, and using those criteria makes it possible to analyze different models of the software development. While, the previously mentioned approaches tend to specify a software development process (i.e. one possible model of the software development). Therefore, the value-based software engineering is a more appropriate as a background theory for my work.

In what follows, a solution for the issues identified during the GQM+Strategies pre-studies (Chapter 4) is presented. For the purpose of evaluating the solution, the VBSE framework was used. The specific components of the solution are mapped into the VBSE key activities. The mapping helped evaluate the approach analytically and identify missing elements that need to be added for a complete solution that can be used to manage organizational value creation process.

6.1 Problem Definition

Any organization that attempts to manage the value creation process needs to first define the organizational value proposition in a format that makes it manageable, i.e. they are able to monitor, control, invoke and implement corrective actions. In other words, the
organizations have to solve the following two-sided problem, or the organizational value management problem. On the one side of that problem is the concept of value and the question: what is valuable for the organization? This is referred to as the value management challenge. On the other side, an organization is often confronted with the question: how do we define the "right" business and organizational goals? This is the business and organizational goals definition and management challenge. The organizational value management problem is a merger of these two challenges. First, it is hard to define good organizational goals if the value is not understood, i.e. what is valuable for the business is not clear. Second, it is even more challenging to manage value creation at the organizational level if the business and organizational goals are not clearly defined and shared.

These two challenges overlap. In the following sections how those two challenges relate and overlap will be explained, and three subproblems that will be addressed in this chapter are formulated.

6.1.1 Subproblem 1: Enunciating Organizational Value Proposition

The first issue to address is whether the GQM+ Strategies approach can be used to document an organizational value proposition in a systematic way. In particular, the focus is on explaining if and how the GQM+ Strategies grid structure can sufficiently capture organizational value.

Figure 23 depicts the basic idea of how organizational value is defined and managed using a hierarchy of organizational goals. At the left side, is the concept of value. The concept of value that organizations and projects are dealing with is time dependent. In other words, organizations make predictions and anticipate that a certain course of actions will result in a value realization for the organization. Such value definitions are subjective to some extent, because different groups of people within organization, e.g., business owners, managers, product owners, developers, etc., can give different answers to the following related questions:

1. What is valuable for the business;
2. What is valuable for the shareholders;
3. What is valuable for the clients; and
4. What is valuable for the employees?
Answers to these questions can be given only by people who are affiliated with or work for the organization. Those people must be able to envision and perceive what is valuable for the organization and formulate that as a value proposition.

![The Concept of Value](image)

**The Concept of Value**

What is valuable for the business?
What is valuable for the stakeholders?
What is valuable for the clients?
What is valuable for the employees?

1. **Value Perception**
2. **Value Expression**
3. **Goals Achievement**
4. **Value Realization?**

**Fig 23. The illustration of the basic idea of how value is perceived and expressed in a goal format.**

The value proposition represents the stakeholders’ (traditional management literature refers to them as stakeholders\(^{36}\)) beliefs and understanding of strategies what in a current situation could be a lucrative option to take. For example, stakeholders may perceive that including a certain type of product in an organizational portfolio is valuable for the business. As a result of that perception, i.e. value proposition, a set of organizational goals can be defined, such as product development goals, new market positioning goals, etc. Traditionally, stakeholders were seen as a "separate" community, usually a numerical minority that did not include the rest of the employee population. Such a distinction has negative impacts on productivity, creativity and, consequently, value creation in knowledge-creating companies like those in the software industry. Therefore, it is necessary to expand the old stakeholder concept to include the creative workers within the organization. To accommodate this change, these two communities must work together. Business owners and management elites (the stakeholders in the traditional view) need to understand why and how the creativity of employees

\(^{36}\)Edward Freeman introduced the term *stakeholder* in his Stakeholder Theory in 1980s [ref]

142
contributes to organizational value. At the same time, it is important that employees become comfortable with the new concept of stakeholders.

The value perception and formulation of value propositions are tightly connected with organizational performance management, i.e. the ability to steer an organization in a direction that will realize the desired value propositions. Therefore, it is convenient to express those value propositions in the form of organizational goals. Furthermore, organizations are complex socio-organizational systems meaning that a unique organizational value proposition represents a whole collection of synchronized value propositions from different groups and units within an organization. If a single value proposition is expressed with an organizational goal then an organizational value proposition is defined through a hierarchy of organizational goals.

The hierarchy of goals (Figure 23) is used to steer and guide operational-level activities in organizations. The information about goal achievement is fed back to the goal owners as only those people are capable of evaluating if the value realization is as they perceived in the beginning. According to the value realization evaluation, further corrective actions can be initiated.

There is a clear connection between these two problems the organizational value creation management and defining and managing the organizational hierarchy of multi-level goals, and solving the latter problem can help with addressing the former. Therefore, the problem can be formulated as how to define the hierarchy of organizational goals that have following characteristics:

– The goals are based on stakeholders’ perception of value;
– They can be instrumented and used for managing organizational activities;
– They are measurable, i.e. it is possible to attach measurement schemes to track organizational progress; and
– They can feed back relevant information (different forms of analysis) about goal achievement to the people (stakeholders) in an organization.

Furthermore, such a goal structure can be used to perform cost and benefit analysis of the organizational goals and strategies, which leads to the second subproblem.
6.1.2 **Subproblem 2: Integrate Cost and Benefits Estimates of Value Propositions**

Presuming that the GQM+ Strategies grid structure can capture organizational value, the second step is to provide mechanism to integrate cost and benefits estimates for the organizational value proposition. This financial-like quantification helps to better understand an organizational value proposition.

6.1.3 **Subproblem 3: Value Realization Risk**

As discussed above, an organizational value proposition is a result of stakeholders’ perceptions and expectations. Due to the inherent uncertainty of the business environment, such perceptions carry a certain level of risk, such as risks of failing to achieve organizational goals, which consequently, leads to value realization failure. An organizational value proposition is a synchronized collection of value propositions across an organization represented by organizational goals. Therefore, it is important to establish mechanism for managing risks that concern the realization of goals.

The existing literature on the GQM+ Strategies do discuss the process of defining the hierarchy of goals (section 2.3). However, there are no contributions that explain in details how to formalize goals using the goal template, i.e. how to refine an informal goal statement into a value proposition documented with the GQM+ Strategies goal template. Furthermore, integrating the cost-benefits and risks into the grid is a completely new feature added to the GQM+ Strategies.

6.2 **Solution Part I: Enunciating Organizational Value Proposition**

This section will discuss how the GQM+ Strategies grid derivation process produces an organizational goal hierarchy that meets all the criteria set by the first subproblem, including a detailed explanation of how to formalize value proposition statements with the GQM+ Strategies goal template.
6.2.1 Defining the Hierarchy of Goals with the Grid Derivation Process

The GQM+Strategies grid derivation process (Section 2.3.2) can start at any organizational level and progressively expand to involve other organizational levels and units. The main output of the process is the grid structure, and an organizational goal hierarchy is contained in that grid structure.

During the grid derivation process, people from various levels define their own goals (goal owners) and strategies to be linked with upper-level organizational goals. By doing this, they are building hierarchy of organizational goals.

The whole process is iterative and involves people, or stakeholders, from various organizational levels. The derivation process involves people in a cooperative and constructive way, by asking them to define their own goals and strategies that will support upper-level goals and strategies. This form of flexibility allows people at different levels to define goals and strategies within their comfort zones that are constrained by upper-level goals and strategies. In other words, people at different levels are provided with bounded flexibility that allows them to optimize locally while they are aligning their activities with global organizational goals and strategies.

Bounded flexibility is the mechanism that allows people to "look around," perceive the value, and think about the right course of action. So, goals defined during the grid development are the result of the goal owners’ perceptions of value.

6.2.2 Formalizing the Value Propositions with the GQM+Strategies Goal Template

Business and organizational goals are formalized using the GQM+Strategies goal template (Section 2.3). The template documents the basic activity that should be performed in order to accomplish the goal, the main (quality) focus of the goal, the object under consideration, the "size" of a change. i.e. the magnitude, the timeframe in which the magnitude has to be achieved, the scope, and basic constraints that may limit accomplishing the goal. Furthermore, potential relationships with other goals are listed (an example of a goal is given in Table 12).

During the grid derivation process, a hierarchy of goal statements is identified. An example goal statement might be: improve customer satisfaction. Goal statements
are evolved and documented with the goal template. Formalizing goals with the template is a process of specifying goals in a qualitative and quantitative manner with a twofold purpose: clarify goals by specifying the dimensions and define the criteria for determining when the goal owners believe the goal can be achieved (magnitude and timeframe).

In the following section goal statement formalization using the GQM*Strategies template will be discussed.

**The Focus and Object of a Goal** Begin by thinking about the focus or the object first because it represents the essence of the goal. The *focus* defines a phenomenon of concern, i.e. what is the attribute the stakeholder cares about. Each focus should be elaborated with an operational definition. The operational definition documents (and makes visible) an organization’s common understanding of what the focus is.

Here are the examples of what a focus can be:

- **Process related**: effectiveness, cost, conformity, etc.;
- **Product related**: quality, usability, defect-freeness, etc.;
- **Project related**: cost, schedule, etc.; or
- **Business related**: updated competencies, employee satisfaction, etc.

The *object* is what should be looked at to deal with the attribute defined by the focus; it directs attention to a cause and specifies the source of changes that must take place. The operational definition of focus can be used as a hint as to where to look for a change.

An object can be:

- **Process related**: process model, sub-process, process area, specific practices, etc.;
- **Product/service related**: end product, intermediate products/artifacts, etc.;
- **Project related**: specific types of project(s), involves process, product, and business component, etc.; or
- **Business related**: organizational, market related, human related, etc.

The combination of the object and focus bounds the activity and acts as a main place to concentrate attention; e.g. improve the quality of the delivered product and evaluate the effectiveness of the process in achieving that product.

For example, for the goal statement "improve customer satisfaction," the focus is *customer satisfaction* and the object is the *delivered product*. To make the goal more specific, it can be specified as customer satisfaction of the delivered product, where
the operational definition of the focus specifies that the customer satisfaction is the
clients’ satisfaction with the quality of the software, i.e. defect free and easy to use.
Furthermore, the object is clarified as the source code and the running product.

**The Scope of a Goal**  The scope of a goal defines the goal’s responsibility zone; i.e.
what part of the organization is responsible for achieving the focus on the object. The
scope defines the goal’s habitat more specifically.

The following are some examples of a scope:

- Entire corporation;
- Certain organizational units;
- Certain projects; or
- Teams or individuals.

For an example of the customer satisfaction, the scope may be *product development
teams located in India*.

**The Activity, Magnitude, and Timeframe of a Goal**  The activity indicates the type
of action associated with the desired change. For the example of customer satisfaction,
the activity is *to improve*; where, at some lower-level a related goal can have for the
activity *to decrease*, e.g., defect slippage.

In general, two categories of activities were identified. The first one is plan to change
or maintain an existing trend, which is “directional” activity, such as decrease, maintain,
increase, and improve. The second category of activities is “developmental” or actions
that are aimed at creating something new that did not exist before. Such activities can
be: to develop, establish, modernize, and build.

The magnitude specifies an expected "amount" of a change for goals with directional
activities and the characteristics that must be met to accomplish goals with developmental
activities. Magnitude can be specified in absolute (AU) or relative (RU) units. Relative
units need to be converted to absolute units later. When specifying relative units, a clear
baseline is necessary. For the customer satisfaction example, the magnitude is specified
with a relative unit of *50% fewer customer complaints* from the average number of
customer complains in the past 12 months (the baseline).

The timeframe specifies a moment when a goal is expected to be achieved, i.e. it
bounds the goals’ existence in time. Timeframe can also be specified in absolute or
relative units. Absolute units represent the calendar time.
The Constraints and Related Goals. Constraints impact and limit a goal’s success. Constraints represent characteristics of the business environment and the organization itself that reduce the goal’s maneuvering space, like identifying the organization’s change or impact limits, e.g. changing them would be too expensive. An example of such a constraint is the global economical crisis.

Table 12. Organizational goal formalized with the GQM*Strategies goal template for an example of improving customer satisfaction.

<table>
<thead>
<tr>
<th>Goal Statement</th>
<th>Improving customer satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal Dimensions</td>
<td>Formalized Goal</td>
</tr>
<tr>
<td>Activity</td>
<td>Increase</td>
</tr>
<tr>
<td>Focus</td>
<td>Customer satisfaction</td>
</tr>
<tr>
<td></td>
<td>(the quality attributes: defect free and easy to use)</td>
</tr>
<tr>
<td>Object</td>
<td>Delivered product</td>
</tr>
<tr>
<td></td>
<td>(the source code and running product)</td>
</tr>
<tr>
<td>Magnitude</td>
<td>50% less customer complains than 12 months average</td>
</tr>
<tr>
<td>Timeframe</td>
<td>6 months</td>
</tr>
<tr>
<td>Scope</td>
<td>Development teams in India</td>
</tr>
<tr>
<td>Constraints</td>
<td>Technology change, etc.</td>
</tr>
<tr>
<td>Relations</td>
<td>CMMI SPI goals, etc.</td>
</tr>
</tbody>
</table>

Relations are related goals that might interact or influence the achievement of the goal. For example, a related goal can require the same resources (e.g. budget), and therefore, it can affect the successful achievement of the goal.

Finally, the goal statement is formalized with the GQM*Strategies goal template, given in Table 12. Goals formalized in this way are further used to:

- Identify what should be measured and define the GQM measurement schemes;
- Identify where the metrics data should be collected as indicated by the scope of a goal;
- Set up the measurement baselines that can interpret the focus, object, and magnitude of a goal.

When the GQM*Strategies goal hierarchy is developed in cooperation with all the relevant factors in an organization, i.e. stakeholders, it enunciates an organizational
value proposition. Furthermore, such value proposition, when formalized with the GQM*Strategies goal template, is measurable, which is an important condition for successfully managing value realization.

6.3 Solution Part II: Integrate Cost and Benefits Estimates of Value Propositions

In the previous section, how the GQM*Strategies grid can be instrumented and used for enunciating an organizational value proposition, as represented by a goal hierarchy was explained. In this section, the focus will be on integrating the goal and strategy cost and benefits estimates using cost–benefit analysis.

The constituents of business value infiltrate the entire organizational structure; therefore, it is difficult to understand the true business value of an organization (Section 2.3). However, the business value should be the most visible at the highest level of an organization for which the business value is important. The GQM*Strategies method provides a structure and process for deriving the goals in a given organizational context, while actual data are used in the GQM graph to interpret the success of realizing organizational goals and strategies. In order to analyze the value aspect of business and organizational goals, it is necessary to implement mechanisms for integrating cost and benefit estimates from different organizational levels, deal with different timeframes of different goals, and structure the estimates according to the recursive nature of the GQM*Strategies grid.

The organizational value analysis is a continuous process based on predictions and assumptions at the beginning with the capability of using actual data during the execution phase. Such capability is essential for understanding how real-world situations impact the value of organizational goals.

The GQM*Strategies grid derivation process is flexible as to the starting point for deriving the grid. The value analysis approach presented here does not limit that flexibility. However, to simplify the presented material without forfeiting generality, the following presuppositions have been made:

- **Presupposition 1.** The GQM*Strategies grid derivation process is top-down, starting with top-level business goals;
- **Presupposition 2.** The value analysis steps are parallel to the grid derivation process;
Presupposition 3. The entire process of the grid derivation and business value analysis is done by a single group with knowledgeable representatives from all organizational levels.

Constraining the inherent flexibility of the grid derivation process has no impact on the methods used; the value analysis could be done after the entire GQM+Strategies grid is defined. However, a parallel derivation can increase the quality of the strategic decisions using goal value-related information to direct the derivation process.

Example. Throughout the dissertation, an example grid for a hypothetical company, BigCorp, is used. The whole example is given in Appendix 2, while the pieces of that example are used in the dissertation. BigCorp is an international company in the business of embedded systems. The main products are robust, mobile software applications developed on the BigCorp’s own hardware platform. The concept of manufacturing hardware and software is attractive for specialized industries, e.g. mining or military. The top-level goal is to increase net earnings 160% in five years (Table 13).

Table 13. BigCorp’s top-level organizational goal.

<table>
<thead>
<tr>
<th>Goal Statement:</th>
<th>Increase net earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal Dimensions</td>
<td>Formalized Goal</td>
</tr>
<tr>
<td>Activity</td>
<td>Increase</td>
</tr>
<tr>
<td>Focus</td>
<td>Net earnings</td>
</tr>
<tr>
<td></td>
<td>(revenue after taxes)</td>
</tr>
<tr>
<td>Object</td>
<td>All product lines</td>
</tr>
<tr>
<td></td>
<td>(software products and services)</td>
</tr>
<tr>
<td>Magnitude</td>
<td>160%</td>
</tr>
<tr>
<td>Timeframe</td>
<td>5 years</td>
</tr>
<tr>
<td>Scope</td>
<td>BigCorp, All units</td>
</tr>
<tr>
<td>Constraints</td>
<td>Global crisis</td>
</tr>
</tbody>
</table>
| Relations       | -/-

A set of strategies for the top-level goal include: increasing sales, improving the production processes, adopting agile practices, implement key agile practices, reducing manufacturing costs, decreasing process variation, and increasing the responsiveness of customer support.
6.3.1 The Value Dimensions of the GQM*Strategies Goal

In order to incorporate cost and benefit estimates into the grid, the goal template with additional goal dimensions, referred to as value dimensions, must be amended. The introduction of the value dimensions will enable the effective integration of the analysis of organizational value propositions represented by the GQM*Strategies grid.

Table 14. The value dimensions of the GQM*Strategies goal.

<table>
<thead>
<tr>
<th>Value dimension</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of goal</td>
<td>Goal specific costs, i.e. the costs directly associated with the goal</td>
</tr>
<tr>
<td>Inherited cost</td>
<td>Total costs of the derived goals. Aggregated costs of lower level goals connected with the goal</td>
</tr>
<tr>
<td>Estimated goal benefits</td>
<td>Expected benefits of achieving the goal, i.e. the benefits directly associated with the goal</td>
</tr>
<tr>
<td>Inherited benefits</td>
<td>Total benefits of the derived goals. Aggregated benefits of lower level goals connected with the goal</td>
</tr>
<tr>
<td>Time period of analysis</td>
<td>A period of time for which the cost and benefit analysis is done</td>
</tr>
</tbody>
</table>

The value dimensions support the evaluation or analysis of the value of organizational goals at certain points in time. In order to analyze value, it is necessary to appraise or estimate all costs and benefits. The main input components, costs and benefits, are time dependent. Without a predefined time period, the evaluation is meaningless. Therefore, a time period of analysis is defined as a separate dimension (Table 14). The time period of analysis is the same for all the goals in a given grid.

There are constraints that have to be considered during the analysis. The most common constraints come from existing contracts (external) or obligations defined by organizational policy or other corporate plans (internal). All those constraints are documented as additional context and assumption elements.

6.3.2 Integrating Cost Estimates

In order to analyze and estimate costs of organizational goals, the recursive structure of the costs must be dealt with. Business or organizational goals represent desired future states while strategies represent means for achieving the goals. In that sense, actual cost carriers are actions that will lead to a desired future state. Further on, the cost of the
strategy can be analyzed using the costs of derived goals, and so on. The recursion is stopped when the goal derivation process reaches the operational level. The allocation of resources is mainly done at the operational level. The following two value dimensions help integrate cost figures and deal with the recursive nature of costs.

The ideal units for expressing costs (and later benefits) are monetary, in terms of currency or money. However, the use of equivalent units is possible and often much easier than monetary units. For example, a use of the FTE (full time employee) units eases the process of estimating necessary efforts for R&D or software development-based organization. The only condition for the use of equivalent units is that there is a way to convert such units into monetary values (e.g. an organization knows the real cost of the FTE).

Cost of goal Specific costs that are entitled exclusively to a goal are accounted for as cost of goal. Some goals (e.g. top-level goals) might not have associated costs of their own while, specific lower-level goals will have major cost factors. The cost of a goal includes the entire time period of the analysis as specified in the goal dimension named time period of analysis. For some goals, the goal timeframe will be shorter than the time period of analysis; this situation is explained below under goal maintenance costs.

Inherited cost The sum of the costs of the connected next-level derived goals represents the inherited cost. Exceptions are bottom-level goals (leaf goals), which do not have derived goals and whose inherited costs are zero. In cases when a single goal is shared by two or more upper-level goals, its costs have to be divided among related upper-level goals. This is discussed below under shared costs.

Example. For the BigCorp example (Appendix 2), organizational goal $G_4$ of establishing an agile software development process inherits €366k in costs from lower-level goals: $G_6$, $G_7$, and $G_8$, i.e. the costs for establishing practices of test-driven development, collective code ownership, and continuous refactoring. The $G_4$ adds €80k of direct costs that are allocated to the additional training and follow-up meetings for the period of five years. These are the total costs that will be allocated to BigCorp’s R&D unit for implementing new organizational strategies to improve the production processes and adopt agile practices.
Goal Maintenance Costs

At the moment when an organizational goal is achieved, some resources should be allocated to maintain the desired state of the goal, i.e. goal maintenance costs. If a goal is in the maintenance state it means that the goal is achieved and actions are performed to maintain the achieved status.

A GQM* Strategies grid contains organizational goals with different timeframes. For example, the derived goals have equal or shorter timeframes from the upper-level goals. In other words, from the cost perspective, achieved goals will be incorporated into the cost analysis with their goal maintenance costs.

Example. BigCorp goals that have shorter timeframes than the time period of analysis, i.e. five years (Appendix 2), are analyzed with respect to maintenance costs. The total costs for increasing market awareness about BigCorp’s products and services ($G_3$) for the five-year period is €350k, from which 250k is for the first two years (the costs for achieving goal $G_3$), and the rest is for maintaining the market awareness in remaining three years. Usually, the goal maintenance costs include some royalties or licenses for software development tools, like for goals $G_7$ and $G_8$.

Shared Costs

Shared costs occur in situations when a goal is connected with several upper-level goals, i.e. when the achievement of a goal contributes positively to a collection of the upper-level goals. Therefore, cost sharing indices for these upper-level goals need to be defined. The sharing index defines the portion of costs (or benefits) that a goal inherits from a lower-level goal. For example, a goal $G_x$ contributes to goals $G_i$, where $i = 1, 2, ..., n$ and $n$ is a number of upper-level goals to whom goal $G_x$ is connected (through some strategies). The designers of the grid have to define sharing indices $\kappa_1, \kappa_2, ..., \kappa_n$, where,

$$\frac{\kappa_i}{\sum_{j=1}^{n} \kappa_j} \times 100\%$$

represents a percentage of how much goal $G_x$ contributes to a goal $G_i$.

Ways to define sharing indices is a discussion for itself. The simplest way to define sharing indices is to give an additional task to the experts so they can define the percentage the cost of a goal will be shared by upper-level goals.
Example. The goal of improving the customer interaction process ($G_{10}$) contributes equally to the goal of increasing sales, especially to the existing clients, and to the goal of improving process performance (goals $G_1$ and $G_2$). Therefore, the sharing ratio is 50%, and the costs and benefits are equally shared among upper-level goals (Appendix 2), i.e. $\kappa_1 = \kappa_2 = 0.5$ for the goal $G_{10}$.

6.3.3 Taking Benefits into Account

Unlike costs, benefits can be harder to estimate and quantify. One of the reasons is that main costs are allocated by lower-level goals (e.g. bottom level) in a grid, and such goals are specific enough that it makes costing them much easier than costing some top-level goals. With benefits, the situation is inverse, i.e. the major benefits are expected and easy to see from higher-level (e.g. top level) goals, and usually, behind the definition of these goals stand long-term predictions and assumptions, which are not as certain and specific as for lower-level goals.

 Nevertheless, it is important to analyze the benefits at all levels (e.g., a benefit at the top level can be expressed as increased market share, while a benefit at the operational level for the same top-level business goal can be effort savings). The following two value dimensions help integrate benefit estimates and deal with the recursive nature of benefits.

**Estimated goal benefits** This dimension represents the benefits that are associated directly with a goal. Sometimes, benefits for lower-level goals might not be as obvious as for top-level goals. The most typical benefits at lower levels are cost saving and resource optimization.

Example. For the BigCorp goal of establishing collective code ownership ($G_6$) it was obvious that some "redundant" activities, such as paired programming and group sessions, will create extra costs, but it was not so obvious that one advantage of those activities was a significant reduction of the costs of familiarizing new team members (assumption A5*).

**Inherited benefits** This value dimension represents a goal’s total benefits inherited from the next lower-level connected goals that are refined from the goal. Similar to costs, benefits are propagated through the goal hierarchy, but unlike costs, the goal
owners (especially at the top level) have clear ideas about the expected goal benefits. The analysis of the grid helps to identify if those expectations are realistic.

**Example.** The BigCorp top-level goal is to increase net earnings for 160% in five years ($G_0$). After analyzing the benefits of derived goals, the expected benefits of the top-level goal are about 85% of the original target. Based on that information, the management board can correct the original target of 160% to 136% or decide to look for additional strategies that can deliver more earnings.

The upper-level goals inherit the benefits of the lower-level goals, and when some goals are shared, additional attention is required, as described below.

**Shared Benefits and Conflicting Goals**

In a situation when a goal is shared by several upper-level goals, the same logic as for the shared costs applies for the shared benefits. This means that, the benefits of a shared goal will be split with the goal sharing indices among the upper-level goals.

However, benefits can manifest the mutual cancellation, which is not a case with costs. This is possible when some goals, e.g. contradicting goals, compete for the same benefits in a mutually exclusive way. For example, company ABC is in the consulting business. Their service delivery channel depends on a software product developed in-house. In order to cope with new technologies, the company established a new department responsible for delivering the services via a global network. A new web-based version of the existing software has been developed. Old and new departments can have the same business goal to increase the market share, but in certain market segments they are targeting the same pool of clients. If the targeted market segments are not overlapping, then there is no impact and no cancellation of the benefits. However, if they do overlap, then it is necessary to analyze and estimate the size of shared benefits among several goals. The number of conflicting goals is expected to be small, so this kind of analysis is feasible. Therefore, it is important to analyze the whole grid structure and not only individual goals.

**Example.** In Appendix 2, the results of cost–benefit analysis for the BigCorp’s grid are given. For the selected strategies, a rough estimate of the return on investment is 10:1, which is an expected return for BigCorp’s business.
6.4 Solution Part III: Value Realization Risk

Due to the inherent uncertainty that surrounds organizations there are no guarantees that organizational goals will be achieved. Failing to achieve an organizational goal, consequently, threatens the value realization of the organization. The possibility of the failure to realize value as perceived by stakeholders is the value realization risk, and that risk can be associated with failing to implement organizational strategies and goals. Therefore, it is important to have mechanism to identify risky goals in the GQM+Strategies grid.

6.4.1 Risk-Related Dimensions of the GQM+Strategies Goal

In order to deal with value realization risk, it is necessary to attach the risk-related information to GQM+Strategies goals. The additional dimensions of the goal template are recommended in Table 15.

<table>
<thead>
<tr>
<th>Risk dimensions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARE</td>
<td>Acceptable risk exposure</td>
</tr>
<tr>
<td>RE</td>
<td>Goal's risk exposure</td>
</tr>
<tr>
<td>Key assumptions</td>
<td>Key assumption that impact goals realization</td>
</tr>
</tbody>
</table>

The focus here is not on a risk analysis method per se, but rather, on integrating the risk-related information into the GQM+Strategies grid. Therefore, this approach is not dependent on a risk quantification method. It only requires that all participants are familiar with the risk quantification method used and its relationship to the acceptable risk model. Some advanced methods for quantifying risk can be found in [72, 126, 173].

Two different perceptions of the risk are used to identify risky goals. The first is an investment-like perception, and it gives a model of an acceptable risk. The second is the risk concerning the realization of the goal itself. By intersecting these two perceptions, the risky goals in the grid are exposed.
Acceptable Risk Model

The rationale for investment-related decisions [172, 173] is based on a minimum of three components: size of the investment, the risk, and the expected ROI. Knowing the level of the financial resources available and the ROI model makes it possible to assess the acceptable risk level.

Acceptable risk is the amount of risk that business owners (investors) are willing to take on in order to materialize perceived benefits in a certain period of time. The acceptable risk exposure (ARE) is a measure of that acceptable risk:

\[
ARE = ARE(I, B(I), t),
\]

where \(I\) is investment size, \(B\) is relative size of benefit from investment \(I\), and \(t\) represents the time period in which the investment is supposed to be realized.

The ARE model is defined for the GQM+ Strategies grid and used to specify the acceptable risk exposure for each goal in a grid based on the cost–benefits estimates of the goal.

Example. The example in Appendix 2 gives an illustration of the ARE model (Equation 37) as a simple 3 × 3 matrix of investment size and benefits. Both, investment and benefits sizes are quantified as: 1–small, 2–medium, and 3–large. The investment size \((I)\) is categorized in respect to the absolute number of investment units (money or any equivalent), while the benefits size \((B(I))\) is categorized as the relative quantity of the invested amount. For the given available resources BigCorp management board characterized an investment of less than €0.5 million as small, an investment of greater than €0.5 million and less than €0.8 million as medium, and greater than €0.8 million as large. Where, for example, the small benefit is less than 30% of investment, medium is between 31% and 65%, and the large benefit is if the investment is greater than 65% of investment (for a time period of analysis). For a mid-level goal \(G_{10}\) the estimated costs (investment) is small, and estimated benefits are medium. Using the acceptable risk model for that mid-level goal gives the \(ARE = small \times medium\) is high. Where the following scale of risk exposures was used: \(VH\) is very high, \(H\) is high, \(M\) is medium, \(L\) is low, and \(VL\) is very low.
Goal's Risk Exposure

The assumption elements of the GQM*Strategies grid are carriers of the uncertainties. A properly derived GQM*Strategies grid identifies and documents assumptions that are key to the successful realization of organizational goals. A set of key assumptions, A, is defined for an organizational goal G. The risk of goal G can be assessed by analyzing the level of certainty of the key assumptions (likelihood), and its impact on the goal fulfillment:

\[ RE(G) = L(A) \times I(A,G), \]  

(2)

where \( RE(G) \) is the risk exposure of an organizational goal G, \( L(A) \) is the likelihood of key assumptions A, and \( I(A,G) \) is the impact of A on the realization of goal G.

Positive and negative types of assumptions are differentiated. Positive assumptions predict a desirable situation, while negative assumptions describe a situation in which actualization has a direct impact on a goal’s failure. For a proper risk assessment, it is necessary to negate positive assumptions.

Goals can have several key assumptions, if that the risk assessment is done for each assumption, and the highest risk exposure is taken.

Risky Goal. A goal G is referred as a risky goal if its risk exposure, \( RE(G) \), is higher than the goal’s acceptable risk level, i.e. \( RE(G) > ARE(G) \).

Example. The key assumption for the top-level goal \( G_0 \) is that BigCorp can substitute a more expensive product with their own in the market (A1). This is a positive assumption. Therefore, the experts were asked to assess the likelihood of failing to substitute the expensive products. The assessment was that the likelihood was low and the impact of the assumption on the goal’s success was moderate, i.e. \( RE(G_0) = \text{low} \times \text{moderate} = L \).

6.4.2 Incorporating Risk into the GQM*Strategies Interpretation Model

The risk handling procedure can be divided into two parts. The first part is related to analyzing the uncertainties of key assumptions and quantifying the uncertainties as the risk exposure of a goal. The second part consists of comparing the acceptable risk level...
and risk exposure of organizational goals. This is done during the incorporation of the risk into the interpretation model.

A prerequisite for risk interpretation is that all the value dimensions for all goals in a grid are defined. This includes cost and benefits estimates, as well as specifying the acceptable risk exposure model \((\text{ARE})\) for a grid and analyzing the risk exposures \((\text{RE})\) of goals for a given time period of the analysis.

For each goal \(G\) in a given grid, the interpretation logic is as follows:

\[
\text{IF } \text{ARE}(G) \geq \text{RE}(G) \text{ THEN the risk level of a goal } G \text{ is acceptable; ELSE risk of an organizational goal } G \text{ exceeds the acceptable risk level for that size of investment and ROI, and it needs to be reevaluated.}
\]

It is possible that after the reevaluation, the situation remains unchanged. Those cases have to be discussed during the feedback session, and an explicit decision is required in order to retain risky goals.

By defining the value dimensions of goals, the GQM*Strategies process and tools can be used to determine the context and assumption variables that are relevant for value analysis. The cost structure is specified by the cost related value dimensions. The identification of the risky goals in a grid reveals a critical GQM*Strategies sub-grid.

**Example.** BigCorp’s grid contains two goals that are identified as risky, goals \(G_6\) and \(G_{10}\). The goal of establishing collective code ownership \(G_6\) is perceived as risky \((\text{RE}(G_6) > \text{ARE}(G_6))\) because it depends on the ability of people changing their habits, which was assessed as of high importance for the goal’s success (the likelihood of accepting new practices is medium, but the impact on the goal is significant).

**Critical GQM*Strategies Sub-grid**

The incorporation of the risk level into the interpretation model identifies a *critical GQM*Strategies sub-grid. The critical sub-grid contains risky goals and goals whose realization is directly or indirectly threatened by the risky goals.

The critical GQM*Strategies sub-grid provides valuable information to the managers, i.e., which goals require more detailed analysis and planning, and which goals require more monitoring when the business goals and strategies are implemented.

For all goals belonging to the GQM*Strategies critical sub-grid, it is necessary to reassess the goal’s risk exposure and acceptable risk level. Situations where goal
Fig 24. The illustration of a critical GQM* Strategies sub-grid.

realization is not going according to plan can reveal changes in context factors and assumptions. Such changes have consequences for the goal’s risk assessment.

After the risk reassessment, new goals can become risky while known risky goals can cease to be risky. Actually, risk monitoring is continually changing the configuration of the critical sub-grid. It is expected that in an ideal situation, when prediction capabilities are good and no unexpected events rise, every new critical sub-grid is a sub grid of previous critical sub-grids.

Example. Figure 24 illustrates the critical sub-grid of BigCorp’s grid. Goals $G_6$ and $G_{10}$ are risky goals (shaded goals in the figure). The critical sub-grid is comprised of the risky goals and goals that are connected with them (i.e. goals: $G_4, G_2, G_1, G_0$). The relationships among goals in a critical sub-grid can be interpreted as:

– **The impact of risky goals on other goals.** Risky goals exceed the level of acceptable risk, and failing to achieve those goals will have negative impacts on the connected upper-level goals. For example, failing to achieve $G_{10}$ can cause serious problems to achieving goal $G_1$, and if goal $G_1$ is not achieved, then the top-level goal $G_0$ will not be achieved. A goal failure does not imply total failure of the connected upper-level goals. In other words, the failure will cause deviations from the magnitude (goal target) of the connected upper-level goals. In the case of BigCorp, failing to increase sales 20% CAGR (goal $G_1$) in five years, can cause a smaller increase in net earnings ($G_0$) than originally planned.
The importance of other goals in the critical sub-grid for the risky goals. The risky goals require more resources, i.e. they have high costs, and when compared with the benefits that they will provide in the given time period, the ratio cannot be justified with the goals risk level, meaning that the expected benefits are too low and the risk is too high. Therefore, the risky goals need upper-level goals to justify the investments required for their implementation. For example, goal \( G_6 \) is risky because the costs of establishing collective code ownership are more than three times the expected benefits. The collective code ownership is necessary to establish agile processes (\( G_4 \)). However, the expected benefits of agile processes are high enough to justify the necessary investment in the \( G_4 \). Furthermore, if risky goal \( G_4 \) is achieved, but upper-level goal \( G_6 \) fails, then the R&D unit might end up financing a successful goal \( G_4 \) that will never return the investment because the \( G_6 \) failed.

In general, identifying risky goals requires defining a criterion for deciding which goal’s risk exposures will be considered risky. One possible way is to set the constant value, and goals whose risk exposure is higher than that constant value are risky. However, using the investment-like criterion and defining an acceptable risk model provides additional quality to the interpretation of the critical GQM* Strategies sub-grids, as discussed in the example above.

The advantage of using GQM* Strategies is that it provides an explicit link to the different levels, from the top level to the operational level. This implies that analyzing organizational goals at different levels, helps to analyze benefits and costs at those organizational levels, and the factors of business value within an organization are made visible. Analyzing multi-level organizational goals has multiple benefits. Incorporating the value related information (i.e., cost, benefit, risk) into the grids helps to not only to make better decisions regarding goals and strategies, it also helps plan the resources for implementing those goals and strategies.

6.5 Evaluation of the Approach

Value-based software engineering will be used as a comparative theory for the purpose of evaluating a solution presented in this chapter. If the solution complies with the VBSE principles, it will result in an argument for discussing its capability to capture the organizational value and to provide mechanisms to manage the value realization process. Furthermore, analyzing the extent to which the solution maps onto the VBSE framework
gives valuable feedback about the weak or missing elements, which are used for further improvements of the solution.

6.5.1 Analytical Evaluation: GQM+ Strategies as Notation for the Value-Based Software Engineering

Boehm [38] proposed the value-based software engineering (VBSE) framework to integrate all the aspects of the software creation process into the perspective of value. The VBSE framework is defined with seven key activities, i.e. groups of activities that are required for the successful management of value. Table 16 gives an overview of the mapping between the VBSE and the GQM+ Strategies approach with its extensions.

The stakeholders’ viewpoint is essential for considering several of the key elements of the VBSE. In addition to the reconciliation of stakeholder value propositions, stakeholders’ values are the basis of risk management, monitoring, and change management (Section 2.4). Those key elements are also present during the GQM+ Strategies grid derivation activities while the remaining key elements are present in the grid execution phase, when actual data are collected and fed into a grid.

In the following sections, the applicability of the VBSE key activities to the GQM+ Strategies solution is discussed.

Benefits Realization Analysis

Benefits realization analysis identifies all initiatives that are required to realize potential benefits. Furthermore, it links resources with outcomes of projects and states assumptions about how those outcomes affect benefits realization. Identifying potential benefits is also connected with identification of stakeholders that have interest in or need to be involved with development process. (Section 2.4)

In other words, the outputs of the benefits realization analysis are:

1. Potential benefits identified;
2. Stakeholders identified; and
3. Sets of assumptions about the chain of outcomes and their impact on benefits realization identified.

During the GQM+ Strategies grid derivation process, when an organizational goal hierarchy is defined (Section 6.2.1) the same outputs are provided as for the VBSE
Table 16. Mapping the GQM* Strategies onto the VBSE Framework.

<table>
<thead>
<tr>
<th>Value-Based Software Engineering</th>
<th>GQM* Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefits Realization Analysis</td>
<td>During the GQM* Strategies grid development goal owners define their value propositions and link them with organizational goals and strategies. Furthermore, context and assumption elements are used to capture their predictions about outcomes. (sections 2.3.2 and 6.2.1)</td>
</tr>
<tr>
<td>Stakeholder Value Proposition Elicitation and Reconciliation</td>
<td>Elicitation of business goals and formalizing it with the template (section 6.2.2) is one of the key activities during the grid derivation process. In the latter phases, negotiation may occur after analysis of the conflicting goals or as a part of the feedback session, when all implications of the business goal are analyzed.</td>
</tr>
<tr>
<td>Business Case Analysis</td>
<td>Costs and benefits are estimated while deriving value goals as a part of the GQM* Strategies grid derivation process. If the context allows quantifying benefits and costs, then a calculation of the ROI indicator is possible. (section 6.3)</td>
</tr>
<tr>
<td>Continuous Risk and Opportunity Management</td>
<td>In the execution phase, context/assumption elements are monitored and updated regularly. With each update, acceptable risks and combined risks are re-calculated and, together with actual cost and benefit data, are fed back into the interpretation model. (section 6.4.1)</td>
</tr>
<tr>
<td>Concurrent System and Software Engineering</td>
<td>Not applicable directly, but indirectly the GQM* Strategies helps to better synchronize the concurrent development efforts.</td>
</tr>
<tr>
<td>Value-Based Monitoring and Control</td>
<td>The purpose of the interpretation model is to enable tracking based on actual data.</td>
</tr>
<tr>
<td>Change as Opportunity</td>
<td>The entire concept of modeling business environments with context/assumption elements and monitoring their changes enables recognition of change on time, while the rest of the grid structure provides infrastructure for making decisions about how to use the change in the best possible way. (section 6.4.1)</td>
</tr>
</tbody>
</table>

benefits realization analysis. The rationale for defining goals (the goal statements) includes the potential benefits that an organization or project will gain if the goals
are achieved. The rationales are documented with context and assumption elements, and a goal with such rationales identifies potential benefits. The goal owners are representatives of the stakeholders.

Developing the GQM*Strategies goal hierarchy involves linking goals from different units and levels, and setting assumptions about how they affect each other. Some of those goals can be project-specific goals, representing project outcomes, and they are linked with strategies with higher organizational goals showing the contributions that they bring to those goals, i.e. this is how the chain of outcomes and their impact on benefits realization is incorporated in the GQM*Strategies grid.

**Stakeholder Value Proposition Elicitation and Reconciliation**

The stakeholder value proposition elicitation and reconciliation seeks to identify the value propositions, and if necessary, the opposing value propositions are reconciled (Section 2.4). The main output is a set of synchronized stakeholder value propositions.

Iterating the GQM*Strategies activities for enunciating individual value propositions (Section 6.2.2) and synchronizing them with other organizational goals, i.e. goal hierarchy (Section 6.2.1), results in an organizational value proposition. This organizational value proposition unifies different stakeholders’ value propositions. In other words, it is a form of value proposition elicitation. This process involves different forms of interactions among stakeholders, i.e. goal owners, such as group sessions, brainstorming, and focus groups, and during those interactions, the conflicting goals (i.e. value propositions) are identified and reconciled.

**Business Case Analysis**

The business case analysis represents a cost/benefit analysis of different options. The unquantifiable elements are treated as uncertainties. The main outputs are:

1. Quantification of costs for different options;
2. Quantification of benefits for those options; and
3. Analysis of uncertainties.

If it is convenient, the ROI indicators are calculated for different options and compared to select the most promising one.
The analysis of multi-level organizational goals as described in Sections 6.3 and 6.4 provides costs, benefits, and risk estimates of the GQM+Strategies elements. The information about cost and benefits can be used to calculate the ROI indicator of different strategies, i.e., options. The risk handling procedure is based on analyzing uncertainties that can affect the value realization, i.e., goal achievement.

Continuous Risk and Opportunity Management

Continuous risk and opportunity management tries to identify situations that might cause losses or threaten value realization. As a part of mitigation plans for such situations, new opportunities might emerge. The opportunity management also addresses situations that are in favor of value realization, e.g., the value can be realized much faster or in an efficient way.

The main outputs and elements of the continuous risk and opportunity management are:

1. The characterization of the risky situations,
2. The characterization of the favorable situations,
3. The ability to monitor factors that can lead to those situations; and
4. Ability to act upon the new circumstances.

All the outputs can be obtained with the GQM+Strategies in the following way. The risk handling procedure (Section 6.4.1) helps to identify risky situations that might cause goal failure, and those situations are documented as key assumptions with negative effects on goal realization, i.e., negative key assumptions. In a same way, positive key assumptions are identified, and they characterize favorable situations. Furthermore, during the grid execution the context and assumption elements are monitored to see if those situations are occurring. Based on this information, and actual data about cost, benefits, and risk are re-estimated and re-calculated, which helps identify which goals and strategies should be corrected to better respond to the new circumstances.

Concurrent System and Software Engineering

Concurrent engineering sees organizing development activities as a way to utilize the parallelism of several developmental phases, instead of running them in a strictly predefined order, usually sequential. For example, the Rational Unified Process,
Evolutionary Spiral Process, and other agile processes emphasize the concurrent engineering approach.

This key activity of the VBSE cannot be directly mapped or linked with the GQM+Strategies approach or its extensions.

**Value-Based Monitoring and Control**

Value-based monitoring and control deals with monitoring the realization of the business value of the outcomes at the project and organizational levels. It must be noted that even though a project can be very successful with respect to its cost-related value, it can be a disaster from the organizational value viewpoint (Section 2.4).

This is one of the most challenging key activities, because it requires a definition of a set of metrics that can be used to monitor and control the value realization progress (e.g. earned-value metrics for the projects). The GQM+Strategies is not predefining a specific set of metrics for that purpose, but rather it gives a customizable approach that can identify relevant metrics for each case.

**Change as Opportunity**

Perceiving a change as an opportunity is more about the individual and corporate attitudes and mind sets of people. People tend to resist changes when they do not fully understand the implications of a change on their decisions.

The entire concept of modeling business and organizational environments with the GQM+Strategies grid elements (e.g. context/assumption elements) and monitoring their changes enables recognition of a change in time while the rest of the grid provides infrastructure for analyzing the implications of the change and for making new decisions about how to use the change in the best possible way (section 6.4.1).

The comparative analysis presented here suggests that the GQM+Strategies, and its extensions, comply to a great extent with the VBSE framework. The summary of the mapping will follow. The benefits realization analysis is carried out while analyzing multi-level organizational goals. Assuming that business value is modeled with the GQM+Strategies goals, a process of defining the goals by the goal’s owners represents stakeholder value proposition elicitation and reconciliation. Refining business goals with strategies and documenting relevant context and assumption elements is a
goal-oriented way of conducting business case analysis. With assumption elements as risk carriers, regular context and assumption updates lead to continuous risk and opportunity management. The value-based monitoring and control is supported by the interpretation model. The entire GQM Strategies grid structure enables acting on changes in the best possible way from the business perspective, i.e. viewing the change as an opportunity.

6.5.2 Findings

The VBSE framework is embodied with seven key elements. According to the VBSE, the entire software development cycle is governed by value-based decisions; therefore it is essential to identify successful critical stakeholders and their needs (i.e. value propositions). The GQM Strategies with their extensions could be especially useful in mapping the stakeholders’ value propositions to value-driven decisions. Also, equally important is an understanding of the links among technical decisions, context, and value creation in different situations in order to improve decision making. Here, the GQM Strategies structure helps in better understanding the relationship between context and the value creation process. Documenting the grid elements (goals, strategies, and context/assumptions) captures relevant information about a particular situation and offers an opportunity to study value-based decisions and actions for that situation. Such studies could be a part of the organizational learning process.

The overall conclusion is that the GQM Strategies complies well with all but two key activities of the VBSE, concurrent system and software engineering, and value-based monitoring and control.

– Concurrent system and software engineering. According to the VBSE framework, concurrent system and software engineering (Section 2.4) is primarily focused on the software "production" process. Recently, Boehm [40] discussed how software engineering (research and practice) addresses the three dominant factors (i.e., hardware, software, and human) in different software development approaches. The conclusion is that approaches that favor only one of the factors at the beginning, so called software-first or human-first, are most likely to fail, and only the hardware-software-human concurrent approach can succeed. The GQM Strategies approach is a supportive methodology that helps synchronize the concurrent development phases and balance the human, software, and hardware factors.
Value-based monitoring and control. The problem is that organizations have complex structures that include multiple interdependent projects at different levels. Each of those projects can be tracked with a set of earned value metrics; however, the metrics do not account for, i.e. model, the project interdependencies. Therefore, a further development of the GQM* Strategies is necessary in order to fully accommodate the value-based monitoring and control.

Enabling a full compliance of the GQM* Strategies approach with the VBSE could represent a significant contribution to the VBSE in the form of a goal-driven notation that makes the value decisions explicit and enables the tracking of decisions and underlying rationales.
7 Organizational Earned-Value Metrics

I conceive that the great part of the miseries of mankind are brought upon them by false estimates they have made of the value of things.

by Benjamin Franklin

In Chapter 6, a comparative analysis of the GQM+ Strategies value-based extensions, and the value-based software engineering framework showed that one key area of the VBSE is not supported well, i.e. value-based monitoring and control. Therefore, the focus of this chapter is to evolve the solution in order to address better that key area.

According to the literature (section 2.4) the most common approach used for value-based monitoring and control in software engineering is earned value analysis or earned value management. Earned value analysis is a tool that controls projects’ budgets and schedules during the projects’ executions (Section 2.4.3). Earned value management (EVM) is the result of positive experiences with the Cost/Schedule Control Systems approach that was used in the 1960s and 1970s. However, the EVM does not take into account the stakeholders’ view on value [43]; and quantifying the value of continuous project tasks can be challenging for several reasons: the benefits of completed tasks become visible after project completion and the subjective estimates of the task completion up to date. Furthermore, organizations are far more complex entities than projects. Using the analogy of a project, organizations can be seen as a system of different projects running at different levels and within various units. All these projects have different life spans, and some are interdependent, i.e. the results of one project impact the success of related projects. In this system, organizations need to track actual costs and monitor the value realization, as stated by an organizational value proposition.

The earned value is usually expressed in the monetary equivalent of work accomplished up to a point in time based on the planned (or budgeted) value for that work (Section 2.4.3). The definition captures the essence of the earned value concept that will be used for developing organizational earned-value metrics. Before presenting a solution, the problem and its subproblems are defined in the following section.
7.1 Problem Definition

The basic idea of earned value management is to track progress of the projects by monitoring budget and schedule. This is achieved by comparing the Budgeted Cost of Work Scheduled (BCWS) and Budgeted Cost of Work Performed (BCWP). The approach was originally used to control contracted projects, such as large projects for governmental institutions. Thus, the government’s term for earned value is BCWP. The cost variance is calculated by comparing the earned value to the actual cost. If the actual cost is less than the earned value of a completed task at a certain point in time, it means that the project is achieving its goals under budget. Similarly, if the earned value of a completed task is higher than the planned value of the tasks, the project is ahead of schedule.

Implementing goals and strategies defined by the GQM*Strategies grid can be viewed as multiple organizational projects that have the following characteristics:

– They can be cross-organizational projects, connecting different organizational levels and units;
– They can vary in lifespan, from very short to long-term projects, and such projects need to be synchronized; and
– They can have multiple owners and stakeholders who share interest.

So, the scope of these organizational projects is different than from the classical projects for which the EVM is used.

The main mechanism of the EVM is to use a set of metrics (basic earned value metrics) and indices (performance indices) to monitor the progress of the projects against the project plans. In order to use that same mechanism for monitoring the implementation of the GQM*Strategies grids, the following two subproblems need to be solved.

7.1.1 Subproblem 1: Tracking Actual Costs and Benefits

The process of implementing organizational goals and strategies requires resources (e.g., money, time, people), while the achievement of the goals is expected to gain some benefits. In order to monitor the process, organizations need a mechanism for collecting costs- and benefits-related data while implementing the GQM*Strategies grid.
7.1.2 **Subproblem 2: Organizational Earned-Value Metrics and Indices**

Once, an organization has the following inputs: the cost and benefits estimates, data about actual costs and benefits, and information about goals achievement obtained through the GQM measurement scheme attached to organizational goals. A set of metrics that can effectively monitor and track the progress of implementing the GQM+Strategies grid against its plan (i.e. budget, schedule, and expected benefits) needs to be defined.

The following sections will present solutions for each subproblem.

7.2 **Solution Part I: Tracking Actual Costs and Benefits**

The GQM+Strategies approach is primarily a measurement approach designed to integrate various goal driven measurement initiatives across an organization. Therefore, the collection of the actual costs and benefits will be treated as a measurement problem, and it will be addressed with the GQM approach.

7.2.1 **The Cost–Benefit GQM Graph**

In order to measure actual costs and benefits, a cost–benefit GQM graph have to be defined and incorporated into the grid. The process used for defining the cost–benefit graph is a typical GQM process. The GQM process is described in details by different literature sources (Section 2.2.2), albeit with several differences. First, the assumption and context elements of organizational goals are known and available, along with the value dimensions of goals (costs and benefits estimates) that ease the process of defining metrics. Second, the costs and benefits structure has a built-in recursion that dominates and shapes the entire cost–benefit graph. In other words, a level-i GQM goal collects costs- and benefits- related data for the current level and feeds data to all upper level connected goals.

Each organizational goal (Gi) is linked to at least one GQM goal with the purpose of monitoring and tracking costs and benefits during the execution phase. The form of the GQM goal is given in Table 17.

A distinguishing characteristic of cost–benefit GQM goals is that dimensions of analyze, purpose, and viewpoint are fixed to costs and benefits, monitoring, and business,
Table 17. GQM goal template for measuring the cost and benefits of goals

| GQM goal | Analyze Costs and Benefits for the purpose of Monitoring with respect to Organizational goal Gi point of view Business in the context of Corporation |

respectively. The with respect to dimension is always a corresponding organizational goal (Gi) whose costs and benefits are observed.

Further, the GQM goal is addressed by four questions:

1. What percentage of budgeted (estimated) costs have we spent?
2. Are there any unplanned costs?
3. Are we achieving planned (estimated) benefits?
4. Are there any unplanned benefits?

Two of the questions relate to costs, and other two relate to benefits.

The cost and benefit data are collected in a certain moment in time, either at predefined periods of time or initiated by information needs. Thus, it is important to keep the record of time with collected data. Practically speaking, collecting cost and benefits data is a straightforward process; in the moments of data collection (time), cumulatively assign all the costs (or benefits) for an observed goal that happened until that moment.

How do organizational goals go through different phases or states, i.e. achievement and maintenance (Section 6.3.1). For the purpose of simplifying the cost–benefit metrics notation a simple functional form for the goal’s costs and benefits is introduced.

### 7.2.2 The Notation of the Goal’s Cost and Benefit Metrics

Organizational goals and strategies are main carriers of the costs and expected benefits. Before introducing metrics functions, a simple notation for goals and strategies will be introduced. Let indexed capital letter $G_i$ detonate an organizational goal. If a goal is in the maintenance phase, i.e. it is achieved, it will be write as $G_i^*$. Each goal in a grid, except the bottom-level goals, is addressed with one or more strategies. Strategies
lead to a set of goals that should have a compound effect on the upper-level goal. In other words, strategies for a goal represent the sum of all the effects of the derived goals. Therefore, strategies that are assigned to a goal \( G_i \) will be denoted with the capital Greek letter sigma \( \Sigma(G_i) \) or \( \Sigma^G_i \).

Time will be one of the independent variables in cost and benefits metric functions. The nature of those functions is cumulative, and often the function values, e.g. \( F(x,t) \), will be observed for some specific periods of time, i.e. for an interval \((a,b)\), denoted as:

\[
F(x,t)_{a}^{b} = F(x,b) - F(x,a),
\]

where \( F(x,t) \) is some cumulative function, and \( t \) represent time.

Estimated values will also be differentiated from actual values, such as estimated costs and actual costs. A function with estimated values will be denoted by \( \hat{F}(\cdot) \), while just \( F(\cdot) \) will represent actual values.

**Goal’s Cost Metric Functions \( \hat{C}(G,t) \) and \( C(G,t) \)**

In the previous chapter a simple mechanism to integrate the cost estimates of organizational goals was introduced and explained (Section 6.3). The main idea was to extend the GQM* Strategies goal template with additional dimensions that are related to goal costs. One of the specifics of the grid cost structure is the recursiveness of the propagating costs throughout the grid, meaning that goals beside their specific costs also inherit costs from lower-level goals. Therefore, the estimated cost for a goal \( G_i \) for the period of time \((0,t)\) is formally denoted by the following equations:

\[
\hat{C}(G_i,\tau)_{0}^{t} = \hat{\gamma}(G_i,\tau)_{0}^{T} + \hat{C}(\Sigma^G_i,\tau)_{0}^{T} + \hat{C}(G^*_i,\tau), \quad (3)
\]

\[
\hat{C}(\Sigma^G_i, t) = \sum_j \hat{C}(G_j,t) , \quad (4)
\]

where \( T \) represents the goal’s timeframe, \( \hat{C}(G_i,\tau) \) represents an estimate of the cumulative cost metric function for a goal \( G_i \), \( \hat{\gamma}(G_i,\tau) \) represents the estimate of a goal-specific costs (cumulative until moment \( \tau \)), and \( \hat{C}(\Sigma^G_i,\tau) \) is an estimate of the cumulative costs of all strategies that refine goal \( G_i \), calculated in Equation (4). The Equation (4) makes recursive reference back to the Equation (3). When a goal is
achieved, some resources are allocated for its maintenance, which is denoted by \( \hat{C}(G_i, \tau) \).

Note that maintenance costs for a goal \( G_i \) include maintenance costs for all goals derived from \( G_i \).

Equations (3) and (4) represent a more formal notation of the goal’s value dimensions introduced in section 6.3.1. Observe that the cost of goal defined in Table 14 is presented by \( \hat{\gamma}(G_i, \tau) \) in Equation (3), while inherited costs are represented by \( \hat{C}(\Sigma G_i, \tau) \).

Example. For BigCorp’s (Appendix 2) goal of establishing agile software practices \( G_4 \), the costs of associated strategies and derived goals for the period of five years are \( \hat{C}(\Sigma G_4, \tau) |_{0}^{5y} = 366k \), and the costs directly associated with the top-level goal (e.g. costs of organizing board meetings) for the period of five years are \( \hat{\gamma}(G_0, \tau) |_{0}^{5y} = 50k \).

The purpose of the cost–benefit GQM graph is to collect actual data, including the actual costs of goals and strategies. The cost–benefit graph follows the same structure as the hierarchy of organizational goals, i.e. it measures the specific costs of goals, but it also propagates costs from lower levels to the upper levels. Thus, the metric function for collecting actual costs \( C(\cdot) \) has the same structure as the cost estimate function \( \hat{C}(\cdot) \).

Therefore, the cost metrics function of a goal \( G_i \), is denoted as \( C(G_i, t) \) and can be written with a similar set of equations as the estimated costs function:

\[
C(G_i, t) = \gamma(G_i, t) + C(\Sigma G_i, t) \quad (5)
\]

\[
C(\Sigma G_i, t) = \sum_j C(G_j, t) \quad ,
\]

where the values of the cost metrics are actual values collected through the cost–benefit graph.

**Goal’s Benefit Metric Functions** \( \hat{B}(G, t) \) and \( B(G, t) \)

The goals benefit estimates, as discussed in Chapter 6, share the same structural logic as cost estimates, i.e. they also manifest recursiveness as costs. Therefore, the estimated benefits for a goal \( G_i \) for the period of time \( (0, t) \) can be formally denoted by the following equations:

\[
174
\]
\[
\hat{B}(G_i, \tau) |_{T_0} = \hat{\beta}(G_i, \tau) |_{T_0} + \hat{B}(\Sigma G_i, \tau) |_{T_0} + \hat{B}(G_i^*, \tau) |_{T_0}
\]  
(7)

\[
\hat{B}(\Sigma G_i, t) = \sum_j \hat{B}(G_j, t)
\]  
(8)

where \( T \) represents the goal’s timeframe, \( \hat{B}(G_i, \tau) \) represents the cumulative estimate of the benefits for a goal \( G_i \), \( \hat{\beta}(G_i, \tau) \) represents an estimate of a goal-specific benefits (cumulative until moment \( \tau \)), and \( \hat{B}(\Sigma G_i, \tau) \) is an estimate of the cumulative benefits of all strategies that refine goal \( G_i \), calculated as in Equation (8). Where the Equation (8) makes recursive reference back to the Equation (7). The expected benefits from an achieved goal are denoted by \( \hat{B}(G_i^*, \tau) \). Note that maintenance benefits for a goal \( G_i \) include maintenance benefits for all derived goals from \( G_i \).

Equations (7) and (8) represent a more formal notation of the goal’s benefit-related value dimensions introduced in Section 6.3.1. Observe that estimated goal benefits defined in Table 14 are presented by \( \hat{\beta}(G_i, \tau) \) in Equation (7), while inherited benefits are presented by \( \hat{B}(\Sigma G_i, \tau) \).

\textbf{Example.} For BigCorp’s (Appendix 2) goal of establishing agile software practices \( G_4 \), the estimated benefits of associated strategies and derived goals for the period of five years are \( \hat{B}(\Sigma G_4, \tau) |_{5y} = 500k \). The top-level goal \( G_0 \) has no directly associated benefits \( \hat{\beta}(G_0, \tau) |_{5y} = 0 \) for the period of five years; however, the inherited benefits for the top-level goal represent total benefits realized by all strategies and goals in the grid \( \hat{B}(\Sigma G_0, \tau) |_{5y} = 12320k \).

During the value analysis (Section 6.3), a situation was discussed when conflicting or competing organizational goals can have, as an implication, a mutual cancellation of the benefits. The analysis of such conflicts is important for defining realistic benefits realization plans for goals, i.e., \( \hat{B}(\cdot) \) function for a goal. However, in the grid execution phase, aggregation of the collected actual benefits data is performed as with costs. Therefore, all accumulated benefits \( B(G_i, t) \) for a goal \( G_i \) until moment \( t \) are given by:
\[ B(G_i, t) = \beta(G_i, t) + B(\Sigma G_i, t) \]  
(9)

\[ B(\Sigma G_i, t) = \sum_j B(G_j, t) \]  
,  
(10)

where the values of the benefits function are actual values collected through the cost–benefit graph.

Both metric functions have to be specified over the same metric units. In Section 6.3, it was discussed that costs and benefits can be expressed using monetary units or some other equivalent. Having, both metric functions expressed with same units makes comparison trivial, i.e. comparing costs for a goal vs. benefits for the same goal.

**Example.** For BigCorp goals and strategies the management board made plans for financing the implementation of the organizational strategies. The plan includes spreadsheets for monthly expenditures and expected benefits, i.e. for each goal \( G_i \), the estimated cost function \( \hat{C}(G_i, t) \), and expected benefits realization \( \hat{B}(G_i, t) \) are specified. Figure 25 shows the estimated values for the costs of BigCorp goals during the five-year time period. Each bar in the figure represents the cumulative costs of the top-level goal \( G_0 \) and its associated strategies in different time moments while participation of the
other goals costs is shown by coloring the bars. A similar graph can be plotted for the benefits realization plans for BigCorps goals and strategies.

7.3 Solution Part II: Organizational Earned-Value Metrics and Indices

Previous sections defined all the elements that are necessary to set up an interpretation model based on the earned-value metrics for an organization. In the following sections the basic set of earned value metrics is defined, and later, some derived earned-value metrics are explained.

7.3.1 Basic Organizational Earned-Value Metrics

Earned value analysis is a simple and powerful tool that helps managers analyze the progress of their projects. The analysis focuses on the budget (cost) and schedule by calculating three basic indicators (metrics): BCWS, ACWP, and BCWP (Table 18).

The GQM+Strategies grid with the cost–benefit graph enables tracking of not only cost-related metrics, but also benefit-related metrics, which allows for an extension of the basic set of earned value metrics with PBRS, ABRM, and PBRM (Table 18).

Table 18. The extended set of basic earned value metrics with benefits related metrics.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCWS</td>
<td>BUDGETED COST OF WORK SCHEDULE: the total budgeted cost up to the analysis date.</td>
</tr>
<tr>
<td>ACWP</td>
<td>ACTUAL COST OF WORK PERFORMED: this is what it actually cost to accomplish all the work completed as of the analysis date.</td>
</tr>
<tr>
<td>BCWP</td>
<td>BUDGETED COST OF WORK PERFORMED: the cost originally budgeted to accomplish the work that has been completed as of the analysis date. This is the earned value.</td>
</tr>
<tr>
<td>PBRS</td>
<td>PLANNED BENEFITS REALIZATION SCHEDULE: the total planned benefits realization up to the analysis date</td>
</tr>
<tr>
<td>ABRM</td>
<td>ACTUAL BENEFIT REALIZATION MATERIALIZATION: this is what it actually materialized of the planned benefits realization as of the analysis date.</td>
</tr>
<tr>
<td>PBRM</td>
<td>PLANNED BENEFIT REALIZATION MATERIALIZATION: the benefits realization originally planned to materialize by the work that has been completed as of the analysis date.</td>
</tr>
</tbody>
</table>
It is expected that implementation plans for the GQM+Strategies grid will contain the budget and predictions about the benefits realization. This will allow the estimated cost functions $\hat{C}(\cdot)$ for the top-level goals promoting into the budget for implementing organizational goals and strategies while the benefits realization plans are specified with $\hat{B}(\cdot)$ for the top-level goals.

The level of detail of the plans can vary depending on the goal level. For example, lower-level goals can have very detailed plans (the same level of detail as for a project) while top-level goals can specify quarterly or half-year plans. Let us define $\hat{C}(G_i, t)$ as budgeted costs of a goal $G_i$; therefore:

$$BCWS_{G_i}(t) = \hat{C}(G_i, t),$$ (11)

where $BCWS_{G_i}(t)$ is the budgeted cost of a goal $G_i$ up to the moment $t$ based on the cost estimates $\hat{C}(G_i, t)$. This can be read directly from spreadsheets or charts like Figure 25.

Actual costs are collected, i.e., measured, with the cost–benefit GQM graph:

$$ACWP_{G_i}(t) = C(G_i, t),$$ (12)

where $C(G_i, t)$ is the actual costs of a goal $G_i$ up to the moment $t$.

The main difference between classical earned value metrics and these organizational earned-value metrics is how $\hat{BCWP}$ is defined. The purpose of using earned value metrics is to analyze the progress of executing organizational strategies. Therefore, the progress of realizing goals is of interest. Each goal goes through two phases. The first phase is the implementation of strategies to achieve the goal. In the second phase, after the goal is achieved, it has to be maintained. In other words, the objective is to bring goals into the maintenance phase by achieving them first. Resources are spent in both phases. The progress of realizing a goal is measured by a goal realization indicator. The goal realization indicator $\xi(G_i, t)$ is assessed using the GQM graph that measures the achievement of an organizational goal $G_i$ at moment in time $t$.

**Example.** If the increase of BigCorp net earnings is 135% after four years, and for the target 160% (goal $G_0$) then $\xi(G_0, 4y) = 0.84$. For another BigCorp goal, if the actual reduction of variation in manufacturing processes ($G_9$) is $0.1\sigma$ after a year, then the $\xi(G_9, 1y) = 0.33$.

Sometimes the magnitudes of goals are characterized and measured with nominal scales, and in such situations $\xi$ takes on value 0–not achieved and 1–achieved. If an
ordinal scale is used then it is possible to define categorical values, i.e. phases of goal achievement, for tracking the progress. Then the $\xi$ will take predefined values for each phase of goal achievement.

The maintenance activities are not success-critical, meaning that it is sufficient to perform the maintenance activities as planned and a goal will be successfully maintained, i.e. remain in the achieved state. For monitoring if the maintenance activities are taking place, it is sufficient to compare the budgeted costs and actual costs for the maintenance activities, i.e. ACWP and BCWP metrics. While activities that are the result of implementing strategies are success-critical, the progress of those activities is assessed through the GQM measurement scheme attached to organizational goals. In other words, the measurement scheme assesses the progress of realizing a goal or a goal realization indicator, noted as $\xi(G,t)$, where $t$ is the time moment of assessing goal achievement. Once, the goal is achieved ($\xi = 1$) it is maintained, otherwise it takes on values that indicate the percentage of goal achievement. One advantage of the GQM$^+$ Strategies based earned-value metrics is a goal-driven measurement scheme that provides feedback about goal achievement, which allows more objective estimates of the goals realization indicators. In classical EVA, the task completion was estimated by project managers without any insight metrics, and it tended be a more subjective than objective estimate.

Considering all arguments discussed above, the budgeted cost of work performed for a goal is defined as:

$$BCWP_{G_i}(t) = \begin{cases}  
\xi(G_i,t) \cdot \hat{C}(G_i,T_{G_i}), & \xi < 1 \\
\hat{C}(G^*_i,t), & \xi = 1, t > T_{G_i} \\
\hat{C}(G_i,T_{G_i}), & \xi = 1, t \leq T_{G_i}
\end{cases}$$

where $T_{G_i}$ is a timeframe defined by a goal $G_i$, and $\xi(G_i,t), \xi \in [0,1]$ is the goal realization indicator of a goal $G_i$. If the goal is achieved, then $\xi = 1$ (this also means that the goal is in the maintenance phase). When calculating BCWP, the following cases can occur. If a goal is not achieved, then $BCWP$ takes a percentage of the entire planned budget for a goal (Equation 13). The percentage corresponds to the goal’s realization index. The second case occurs when a goal is achieved on time, as specified by the goal’s timeframe, then $BCWP$ will be equal to all costs that were planned until that moment (Equation 14). The third case occurs when a goal $G_i$ is achieved in time, before the defined goal’s timeframe (Equation 15), then $BCWP$ takes the entire planned budget for the goal implementation.
Example. After four years \((t = 4y)\), BigCorp management follows up the progress of the top-level goal. The achieved net earnings are 135\%, i.e. \(\xi(G_0, 4y) = 0.84\), while the budgeted cost of the work performed (the portion of the costs that can be justified with achieved results) is:

\[
BCWP_{G_0}(4y) = \xi(G_0, 4y) \cdot \hat{C}(G_0, T_{G_0}) = 0.84 \cdot 1046k = 878.64k
\]

and the planned budget for that period is: \(BCWS_{G_0}(4y) = \hat{C}(G_0, 4y) = 903k\).

In the same way, the benefit-related earned value metrics: \(PBRS_{G_i}(t)\), \(ABRM_{G_i}(t)\), and \(PBRM_{G_i}(t)\) can be defined:

\[
PBRS_{G_i}(t) = \hat{B}(G_i, t), \quad (16)
\]

where \(PBRS_{G_i}(t)\) is the planned benefits realization schedule of a goal \(G_i\) up to the moment \(t\) based on the benefits estimates \(\hat{B}(G_i, t)\).

\[
ABRM_{G_i}(t) = B(G_i, t), \quad (17)
\]

where \(B(G_i, t)\) is the actual benefits realization of a goal \(G_i\) up to the moment \(t\).

\[
PBRM_{G_i}(t) = \begin{cases} 
\xi(G_i, t) \cdot \hat{B}(G_i, T_{G_i}), & \xi < 1 \\
\hat{B}(G_i, T_{G_i}), & \xi = 1, t > T_{G_i} \\
\hat{B}(G_i), & \xi = 1, t \leq T_{G_i} \end{cases} \quad (18)
\]

where \(PBRM_{G_i}(t)\) is the planned benefit realization materialization, \(T_{G_i}\) is a timeframe defined by a goal \(G_i\), and \(\xi(G_i, t), \xi \in [0, 1]\) is the goal realization indicator of a goal \(G_i\).

When calculating \(PBRM\), the following cases can occur. If a goal is not achieved, then \(PBRM\) takes a percentage of the entire planned benefits for a goal (Equation 18). The percentage that corresponds to the goal’s realization index. The second case occurs when a goal is achieved on time, as planned by the goal’s timeframe. Then \(PBRM\) will be equal to all the benefits that were planned until that moment (Equation 19). The third case occurs when a goal \(G_i\) is achieved in time, before the defined goal’s timeframe (Equation 20). Then \(PBRM\) takes the entire planned benefits for the goal.

Analyzing earned value metrics can help determine if a goal realization is lagging behind (when \(BCWP < BCWS\)), exceeding budgeted costs (when \(BCWP < ACWP\)), or if the materialization of benefits is lower than planned (when \(PBRM < ABRM\)). More
detailed explanations of how to perform the analysis can be found in [43, 50]. In the
following section, the performance indices that are particularly useful for monitoring
organizational goals and strategies will be defined.

7.3.2 Cost and Benefits Performance Indices

The most common way to perform earned value analysis is to calculate derived
metrics from the basic set of earned value metrics. Here, two such metrics that can be
particularly useful are presented. They are the costs performance index (CPI), and when
it is used for benefits-related metrics, it will be called the benefits performance index
(BPI). The performance indices are defined for a goal $G_i$ as:

$$CPI_{G_i}(t) = \frac{BCWP_{G_i}(t)}{ACWP_{G_i}(t)}$$

$$BPI_{G_i}(t) = \frac{PBRM_{G_i}(t)}{ABRM_{G_i}(t)}$$

when $ACWP_{G_i} \neq 0$ and $ABRM_{G_i} \neq 0$. They show the cost and benefit performance of a
goal $G_i$ until moment $t$. If the result is less than 1, you are over budget/plan. If it is
greater than 1, you are under budget/plan (Table 19).

<table>
<thead>
<tr>
<th>VALUE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CPI = 1$</td>
<td>Goal realization costs are according budget.</td>
</tr>
<tr>
<td>$CPI &gt; 1$</td>
<td>Goal realization costs are less than budgeted.</td>
</tr>
<tr>
<td>$CPI &lt; 1$</td>
<td>Goal realization costs exceed budget.</td>
</tr>
<tr>
<td>$BPI = 1$</td>
<td>The benefits realization for a goal is as planned.</td>
</tr>
<tr>
<td>$BPI &gt; 1$</td>
<td>The benefits realization for a goal is less than planned.</td>
</tr>
<tr>
<td>$BPI &lt; 1$</td>
<td>The benefits realization for a goal is higher than planned.</td>
</tr>
</tbody>
</table>

The interpretation of the benefits performance index differs from the costs per-
formance index (Table 19). A small value of $CPI$ indicates a situation where more
resources are spent than budgeted, but if $BPI$ has a small value, it indicates better
materialization of the benefits than planned. If the denominators of the performance
indices ($ACWP$ and $ABRM$) have values close to zero, they will result in high values of
the performance indices. Therefore, it is advised that the interpretation of performance indices is accepted when their values are within the interval \((0, 2)\), meaning that a variation from budget/plan is at most plus/minus 100\%. If the value is out of the interval, it is a good indicator that something unplanned is happening, and further analysis is required.

Using both performance indices helps distinguish between two situations that would be hard to differentiate with only a cost performance index.

- Situation when the spending for a goal realization is over budget, but the benefits realization for the goal is much higher than planned, i.e. \(CPI < 1, PBI < 1\); and
- When the spending for a goal is over budget, and the benefits realization is not occurring or is much less than planned, i.e. \(CPI < 1, PBI > 1\).

Having information about the goal’s benefits realization increases the quality of organizational decisions. For example, in the first situation, managers can consider tolerating costs over budget because of good benefits realization. While, in the second case, tolerating high costs is not an acceptable option.

Costs and benefits manifest recursive behavior, meaning that budgeted costs for a goal at one level include budgeted costs of derived lower-level goals. Similar, but in a less obvious way, benefits from different levels are aggregated. In the literature, examples were found of situations where the measurement of the benefits in terms of their monetary equivalent is difficult, if not impossible. One such example is customer satisfaction. At a certain level, i.e. where the goal of increasing customer satisfaction is defined, it is not adequate to measure it in financial terms. However, if it is at a higher level, i.e. where the goal of increasing profitability is defined with the assumption that increasing customer satisfaction will increase profitability, then it is possible to measure the benefits of increasing customer satisfaction in terms of financial value in the context of an upper-level organizational goal.

The performance indices can provide a good insight into goals realization at certain moments. Furthermore, visualizing performance indices (plotting on time axis) for goals in a grid reveals different patterns and trends, as described in the following section.
7.3.3 Interpreting Organizational Earned-Value Metrics: Analyzing Effectiveness of Strategies

Plotting performance indices for individual goals can be a useful tool for identifying budgets or schedule overruns or to visualize the value realization for specific goals as they happen. Interpreting earned value metrics (Section 2.4.3) provides information about the goals realization progress with respect to schedule, costs, work (the goal achievement), and the benefits realization that are expected if goals achieved.

Example. The follow up of BigCorp’s top-level goal with earned value metrics after four years is given in Table 20. The goal realization index is $\xi(G_0, 4y) = 0.84$. The management board can read the following from the earned value metrics. The

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACWP$_{G_0}(4y)$</td>
<td>1011k</td>
</tr>
<tr>
<td>ABRM$_{G_0}(4y)$</td>
<td>10348k</td>
</tr>
<tr>
<td>BCWS$_{G_0}(4y)$</td>
<td>903k</td>
</tr>
<tr>
<td>PBRS$_{G_0}(4y)$</td>
<td>11000k</td>
</tr>
<tr>
<td>BCWP$_{G_0}(4y)$</td>
<td>878.64k</td>
</tr>
<tr>
<td>PBRM$_{G_0}(4y)$</td>
<td>10348.8k</td>
</tr>
</tbody>
</table>

realization of the goal $G_0$ is falling behind the schedule ($BCWP < BCWS$), and more money was spent than planned for that period ($ACWP > BCWS$). However, the benefits realization is within acceptable limits and according to plan; therefore, despite the slight increase in costs management, the board decides to continue with the original plans.

The earned value metrics provide a mechanism to follow up the progress of individual goals, where the progress and contributions of the subgoals are included through inherited costs and benefits. In other words, looking at the top level goal earned value metrics indicates how well the other goals in the grid perform. If there are some problems with goals realization, then it is necessary to look at the earned value metrics for all goals in a grid, and that can reveal the exact position of the problem in the grid. For that purpose, it is convenient to use earned value indices, i.e. CPI and BPI. These indices, when plotted for the goals in the GQM$^+$ Strategies grid, can reveal situations were strategies are more or less effective. Those situations are identified by one of the following patterns:
1. **Converging pattern**: the charts form a funnel-like shape. When plots for performance indices for connected goals from adjacent levels converge, which it is an indication of an effective strategy;

2. **Diverging pattern**: the charts are frequently crossing over each other and the values are not converging. This is a clear indication of ineffective strategies; or

3. **Gap forming pattern**: the charts are showing a stable gap among some indices. The size of the gap indicates how wrong the assumptions were about the strategies. The smaller the gap is, the more accurate assumptions are.

Each of these patterns is an indication of the effectiveness of organizational strategies. For the purpose of illustrating patterns, a sub-grid of BigCorp’s grid example will be used. The sub-grid contains three goals $G_0$, $G_1$, and $G_3$; where $G_1$ is a parent goal of goal $G_3$, and $G_0$ is parent to $G_1$. For the purpose of this discussion let us assume that goal $G_1$ has no other derived goals than goal $G_3$. In the following section the situations that might occur are plotted and discussed.

**Converging Pattern as an Indicator of Successful Strategies**

Figure 26 depicts a converging pattern when all performance indices converge toward one value, which in an ideal case is the value of 1. The ideal case is when the spending and benefits realization is exactly according to plan.

![Costs Performance Index and Benefits Performance Index](image)

**Fig 26.** An illustration of the converging pattern. The chart shows CPI and BPI for goals $G_0$, $G_1$, and $G_3$ marked with enumerated dots 0, 1, and 3.

A conclusion that can be drawn if all the performance indices converge is that all the strategies for organizational goals are working, providing the results that were expected when the grid was developed. In other words, all the activities connected with goals
are conducted, resources are being used according to plan, the activities support the achievement of goals, and goals are bring the expected benefits.

If convergence occurs toward some value other than 1, then the distance from the 1 can be used to calculate the approximate percentage of how much over or under plan the value realization (for BPI) or budget spending (for CPI) is for the entire grid on the average.

**Diverging Pattern as an Indicator of Ineffective Strategies**

Figure 27 depicts the diverging pattern when the benefit performance indices are diverging and/or crossing over each others. The frequent crossing is an indication that goal realization is more a result of a coincidence than of a planned strategy. Usually, the values are larger than 1, indicating that benefits realization is under estimated while in case of costs the index takes on small values.

![Costs Performance Index](image1)

![Benefits Performance Index](image2)

*Fig 27. An illustration of the diverging pattern. The chart shows CPI and BPI for goals \( G_0, G_1, \) and \( G_3 \) marked with enumerated dots 0, 1, and 3.*

In order to better explain this pattern, let us start from the previous pattern, when all indices of goals converged and formed a funnel-like shape. If only one goal is not achieving its plan target, the associated costs will not be “justified” by the goals and strategies results, and that will cause a smaller value for BCWP due to a smaller goal’s realization index \( \xi \). Because of that the CPI will have smaller values than CPIs of neighboring goals. Furthermore, if the discernment in values shows this trend over time, it can be interpreted as an ineffective strategy, like CPIs in Figure 27 for goals \( G_0 \) and \( G_1 \). The situation in the figure shows that despite the success of goal \( G_1 \) the goal \( G_0 \) is
not progressing, and this can be explained by an incorrect strategy that connects the two goals.

If a *diverging pattern* occurs, it is a clear indication that other unplanned factors are impacting the realization of the goals. Either the strategies are not effective, or the defined strategies are not sufficient, i.e. there are some missing strategies.

**Gap Forming Pattern as an Indicator of Incorrect Assumptions about Strategies**

Figure 28 depicts the gap pattern when the benefit performance indices between certain levels form a stable gap. The gap indicates that predictions about the contributions of the derived goals were over- or under-estimated.

![Costs Performance Index](image1)

![Benefits Performance Index](image2)

**Fig 28.** An illustration of the gap forming pattern. The chart shows CPI and BPI for goals $G_0$, $G_1$, and $G_3$ marked with enumerated dots 0, 1, and 3.

Furthermore, the size of a gap is an indication of how much the assumptions about a strategy that connects the goals in the gap were wrong (or under/over estimated). For example, if the size of the gap is about 0.3 between goals $G_0$ and $G_1$, then the assumption that achieving $G_1$ will be sufficient to realize $G_0$ is wrong, and the realization of $G_1$ brought only 70% of the planned benefits to $G_0$, or it can be said that a strategy connecting $G_0$ and $G_1$ is 70% effective with respect to planned benefits realization.

The interpretation of the *gap pattern* is significantly different for cost performance indices than for the benefits performance indices. Figure 28 illustrates a gap between CPIs for goals $G_0$ and $G_1$. The size of the gap (e.g. $\approx 0.5$) does not mean that $G_1$ has 50% less costs, it means that 50% of the costs of $G_1$ are not justified. Despite the fact that the realization of $G_1$ goes according to plan (costs for $G_1$ are within the budget),
$G_1$ affects $G_0$ with an impact of 50% of the estimated; therefore, 50% of $G_1$ costs are justified by the upper-level goal $G_0$.

When the GQM* Strategies grid implementation starts, the cost performance indices are more reliable than the benefits performance indices. This is because costs start to flow from the moment the resources are assigned to work, and they contribute to organizational goals while benefits realization happens later, after the successful achievement of the first bottom-level goals. Therefore, it is advisable to use both performance indices.

### 7.4 The Compliance of the Organizational Earned Value Analysis with Value Based Monitoring and Control

The earned value metrics can be used to predict the expected completion date and costs. However, the earned value metrics, when used for monitoring organizational value creation, are an instrument for identifying organizational alignment problems, e.g. when the operations of some organizational units are not aligned with higher organizational goals or the results of the lower-level goals are not producing the desired effects on upper-level goals.

Organizations develop GQM* Strategies grids with the intention to align their goals and strategies and to establish measurement schemes that can provide feedback on goal achievement. When the strategies are defined, in order to implement them, resources are required (e.g. people, equipment, time) and they represent inputs for the GQM* Strategies grid execution process (Figure 29). The expected outputs of that process are the realization of value propositions (e.g. satisfied customers, stronger market presence, and high quality product).

Costs can be used to quantify inputs in terms of monetary value while outputs can be characterized as benefits. For example, test-driven practices can be planned with respect to how much of a resource and how much time must be spent on adopting new practices. At the same time, a successful adoption is expected to decrease defect slippage, which would indirectly decrease the effort spent fixing bugs, and that can be quantified in terms of how much money will be saved with the new practice. The adoption of the practice itself can be measured in terms of process compliance.

The earned value system discussed here is introducing an additional variable that is observed, i.e. benefits, compared to the classical earned value approach where only two
variables are observed: costs and work completion. To explain this, in Figure 29, the possible scenarios are plotted. Actually all the combinations of input, goal realization, and output variables are possible. For instance, if the execution of the GQM*Strategies grid is under resourced, i.e. in sufficient resources are provided, the goal achievement fails, and consequently, the benefits are not realized. The scenario is clear; however, there could be other scenarios that are less clear, and even look contradicting. One such scenario occurs when the goals are achieved as planned, but no benefits are realized. This is possible, with an exception to financial goals for which the benefits realization is directly connected with the magnitude of the goal. For example, BigCorp’s top-level goal $G_0$ of increasing net earnings can be directly converted into monetary units as benefits, while goal $G_6$ of establishing collective code ownership was assumed to bring some savings from familiarizing new team members. The adoption of the collective code ownership can be successful, but no savings are realized.

Another scenario occurs when goals are not achieved, but benefits realization occurs despite goal failure. This can happen when the circumstances change dramatically. For example, a goal can be to become a market leader by overtaking the market from a major competitor. Suddenly, the major competitor drops out due to some dramatic event...
(e.g. earthquake or other hazards). Therefore, to identify all the scenarios it is important to observe all three variables and to include them into the organizational earned value system.

In Chapter 6, the organizational value management problem was discussed, and it explained how people (stakeholders) perceive what is valuable for the business, expressed that as goal hierarchy, and then based on goal achievement that they can reason if the value was realized as they perceived (Figure 23). In order to support this value creation process, the cost related earned value metrics help with implementing organizational goals and strategies while benefits related earned value metrics help analyze if the benefits realization happens as expected while achieving organizational goals.

Value-based monitoring and control (Section 2.4) needs to be implemented for supporting a value realization feedback process (Figure 13). The organizational earned value system presented here (Figure 29) is fully compliant with the value realization feedback process in the following ways:

– Developing business cases, time-phased cost, benefit flows, and plans represents developing organizational goals and strategies, i.e. GQM*Strategies grid and analyzing costs and benefits of the organizational goals and strategies (discussed in Chapters 6 and 7);
– Performance plans are the execution of the GQM*Strategies grid;
– Monitoring if value is being realized is conducted through monitoring goal achievement (cost related earned value metrics) and analyzing the realization of expected benefits (benefit related earned value metrics);
– Assumption validity is checked if the earned-value metrics reveal unwanted patterns, like converging or gap-forming patterns. Furthermore, the costs related metrics can indicate that assumptions regarding strategies, i.e. the effects of lower-level goals on upper-level goals, are valid. The benefit-related metrics indicate if the assumptions that connect goals with expected or perceived value realization are valid. These assumptions are explicitly documented as the GQM*Strategies elements.

The organizational earned value system based on the GQM*Strategies is not just complying with the value-based monitoring and control requirements, it improves the value realization feedback process itself, through the introduction of the benefits-related earned-value metrics.
8 Constructing Organizational Causal Models with the GQM*Strategies

*Shallow men believe in luck or in circumstance. Strong men believe in cause and effect.*

by Ralph Waldo Emerson

In this chapter the dependencies and relationships among goals, i.e. relationships of goal–strategies–goals, will be discussed. Those relationships carry knowledge about organizational alignment. The ability to analyze and quantify them, to some extent, helps package that knowledge so it can be used or reused in different types of queries over the GQM*Strategies* grid. For example, it can be used to query the effect (impact) of an organizational unit’s failure to achieve some lower-level goals on some higher-level goals and strategies.

During the GQM*Strategies* industry pilots, the practitioners suggested that having a mechanism to analyze and quantify the contributions of different goals to upper-level strategies and goals would be a valuable extension of the method (Section 4.4). Having that kind of information can help prioritize activities during the grid execution phase.

The purpose of the GQM*Strategies* interpretation model is to provide an interpretation framework that uses various metric schemes across an organization to contemplate goal achievement and strategy effectiveness. Even for small-sized grids, the number of interdependencies among grid elements is high, and the complexity only increases with larger grids. The relationships and links among grid elements are the source of grid complexity; therefore, it is important to analyze and formalize different types of relationships that can occur during the grid derivation process.

In the following sections, a formalism is introduced to analyze the underlying theory of the relationships among the GQM*Strategies* grid elements. Such analysis can show how the goals in a grid affect each others; for example, how much threat is posed by risky goals over other goals in a grid. Different formalisms may produce different models of the GQM*Strategies* concepts. The formalism used here is based on *causality theory*.

The GQM*Strategies* grid documents organizational or business goals and ways to achieve them, as articulated by strategies at different organizational levels. For example, top-level goals define a future situation that should be an effect of the successful execution of organizational strategies. At the next level, goals are defined with the
intention of having a *causal effect* on the upper-level goals, and so on. Causal thinking is at the basis of the derivation of the GQM* Strategies grid, and the grid itself represents an **organizational causal model**. Therefore, causality theory was the theory selected as a theoretical framework for formalizing the GQM* Strategies concepts.

Defining goals and strategies for achieving the goals is a process of envisioning a sequence of **events** that should cause the desired outcomes. In a generalized form, cause and effect can be defined as events [185]. A common perception of a causal relation is as a mechanism unfolding over time that uses the cause to produce the effect, with one general constraint: effects cannot precede their causes. According to Sloman [185], the perception of a causal relation defined as a mechanism is misleading because such perception implies a condition-like certainty—whenever the causal conditions are fulfilled, the effect should emerge. A more realistic perception involves **probabilistic reasoning**. Therefore, **causal relations relate entities** (events or classes of events) **that exist in and therefore are bound in time** [185, p.22]. A collection of all events form the **event space**.

In business and management literature, the ideas of analyzing current situation and identifying desirable future situations toward which an organization will be steered are most pervasive in the **Positioning School** of strategic management [153]. The strategy development, according to the Positioning School, requires organizations with formalized and centralized structures, preferably in a domain of mass production [153, p.358]. In such organizations, strategy development is a process based on analysis of the facts conducted mainly on the top-level, which results with a smaller number of the top-level goals and sub-goals. There is no need to further decompose or refine those top-level goals, because the organizational structure is so well defined that the role of its organizational units in contributing to those top-level goals is self-evident [153]. Software-intensive or ICT industry do not have such well defined and centralized organizational structures, and often contributions of various organizational units to the top-level goals and strategies is not self-evident. Furthermore, organizational units in order to address some top-level goals need to find creative solutions for the problems posed to them, and those creative solutions need to be made explicit and linked to the contributing top-level goals and strategies. This approach is much more closer to the **Cognitive School** of strategic management than to the Positioning School [153]. Therefore, the idea of developing organization causal models can be seen as a mix of these two schools of strategic management. First, the Positioning School, as the most dominant at the top-management level, identifies positions toward which an organization
needs to go, i.e. top-level goals and strategies. Second, the process of refining those goals and strategies continues through the entire organization utilizing the brain power of people from various organizational units (Cognitive School) to define their sub-goals and strategies, or in terms of Positioning School to define positions of their units that will have an effect on successful positioning of the entire organization.

The concept of events is important because the real world is eventful and inwrought with different kinds of events, and often, individual events are a part of a larger scheme of events, i.e. some events support or inhibit the occurrences of other events. In fact, the whole process of defining GQM$^+$ Strategies goals and strategies aims at perceiving and modeling the favorable events of different organizational units, the events that would lead to a successful achievement of higher organizational goals and strategies.

### 8.1 Problem Definition

This chapter focuses on analyzing the relationships among goals. Answering the following question helps expose certain types of relationships: how likely is the achievement of a certain set of goals (e.g. goals set $Y$) if one set of goals is achieved (e.g. goals set $X_1$) and at the same time another set of goals is not achieved (e.g. goals set $X_2$)? For example, if we define the set $X_2$ as a set of all risky goals and the set $Y$ as the set of all other goals in a grid, leaving the $X_1$ to be an empty set, then the query will give us a probability of having non-risky goals in a grid achieved if all risky goals are not achieved (i.e. the worst case scenario for the risky goals). And, from that query we can shed more light on the relationships among risky and non-risky goals by understanding the magnitude of the impact of risky goals onto other goals in a grid. This type of question, i.e. query over a grid, generates information that is contained in the GQM$^+$ Strategies grids, but it is not obviously visible or easy to extract and access.

In order to use probabilistic reasoning and run different queries over the GQM$^+$ Strategies grid, a probability distribution function is necessary. The probability distribution function contains the knowledge about the dependencies and alignments of different organizational goals, and that knowledge comes from the experts who are building the GQM$^+$ Strategies grid and an organizational causal model. A significant role of the organizational causal model is to help transform expert knowledge that was used to construct the GQM$^+$ Strategies grid into the probability distribution function.

There are different ways to define a probability distribution, because we are dealing with a finite set of goals; therefore, a finite set of outcomes of interest, the probability function
distribution tables (PDT) are used. The PDT is a table-like structure with the fields representing values provided by experts, i.e. probability estimates regarding certain outcomes. Having one “big” PDT for a grid is not feasible, not only because of the number of combinations, but also because of the experts’ potential lack of familiarity with certain parts of a grid (i.e. sub-grids). Fortunately, it is possible to define the probability distribution function with many smaller PDTs interconnected over the dependency lines, as shown in Figure 30.

Figure 30 shows that each PDT corresponds to exactly one goal in a grid, and it defines values concerning possible scenarios for that goal with respect to all the possible outcomes of derived goals. Assessing the information regarding the goal achievement from the experts is also infeasible, with the exception of leaf goals. The experts cannot make a valid claim that a goal in a grid will be achieved with some certainty if some smaller subset of goals is achieved because there are other goals that influence that particular goal. In other words, it is not possible to directly fill in the values for the PDTs in Figure 30 by asking experts. This leads to the problem of how to define probability distribution over the GQM+Strategies grid.

This problem will be broken into two subproblems.
8.1.1 **Subproblem 1: Identifying the right question.**

First, it is important to find the right question, i.e. what kind of information can experts answer with confidence while they are developing the GQM+Strategies grids.

8.1.2 **Subproblem 2: Calculate missing PDT values.**

The second subproblem addresses the issue of generating or calculating missing values in PDTs from the information provided by experts with the help of the relevant theory, i.e. causality theory or providing a sufficient apparatus to define the probability distribution function over the GQM+Strategies grid.

The following two sections offer a solution to the problem. Each section corresponds to one of the subproblems. The first section will identify the right question by elaborating what events are of interest and how those events can be modeled and documented with the GQM+Strategies concepts. Then, causality theory is used as the background theory to formalize the GQM+Strategies concepts in order to develop a sound mechanism for calculating the necessary probability distribution values from the information provided by the experts.

8.2 **Solution Part I: Identifying the Right Question**

Causality theory operates through the concepts of events and event spaces. Those concepts represent or model real-world situations. Sections 8.2.1 and 8.2.2 discuss how the GQM+Strategies concepts of goals, strategies, and context and assumption elements can be used to model events and search the event space through goal refinement patterns. Types of strategies that can be used to analyze any goal refinement pattern will be identified, and a measure of strategy effectiveness will be defined. Refining a single goal with strategies represents a unitary building block of the GQM+Strategies grid derivation process. Understanding that unitary building block of the grid, will help formulate the right question.
8.2.1 **Modeling Events with the GQM*Strategies Concepts**

From the information point of view, an event can be described with a finite set of relevant variables (information carriers). That information is used for decision making. However, for successful decision making the following problem must be addressed—how can the subset of relevant variables be identified? Goals play a significant role in addressing that problem. They represent markers for the information relevance criteria. The concept of relevance is well defined in philosophical texts on contextualism through the principles of indexicality and relevant alternatives (e.g. [68, 77]). For example, a student wants to pass an exam (goal). Knowing the student’s goal, the set of variables describing various textbook should help narrow down the relevant textbooks from which to study. This mechanism shows how goals can index relevant alternatives, i.e. identify the relevant information.

In a similar way, the GQM*Strategies elements can be used for identifying relevant contextual variables and document events.

**Example.** BigCorp’s top-level goal $G_0$ describes a *future event* (timeframe: five years) for the *entire company* (scope) when BigCorp will be making 160% more net earnings (activity, focus, and magnitude) from their *products and services* (object) than they are today. Further clarifications of the event are provided with relevant context and assumption elements (Appendix 2).

The GQM*Strategies goal template models and documents *favorable* events in a structured way. Failing to achieve a goal is also an event, but not a favorable one. In this section, favorable events, or future events are discussed. Later, in Section 8.3.2 a formal definition of an event will be introduced. The goal template defines eight dimensions that need to be specified (more detailed explanation is given in Section 6.2). First, to define an event more clearly an organization needs to specify the goal’s dimensions of activity, focus, object, scope, constraints, and relations. Second, which is very important for the navigation toward a future event, is to define a criterion for determining the degree to which a goal should be achieved—magnitude—and when the goal should be achieved—timeframe. The goal magnitude and timeframe dimensions are interpreted quantitatively.
The use of the GQM*Strategies goal template is discussed in detail in Section 6.2.2. Here, a mathematically formal notation for goals that will be used throughout this chapter, is introduced.

**Goals** Throughout the dissertation, a set of goals $\mathcal{G} := \{G_i \mid i \in N\}$, where $N := \{0, 1, 2, ..., n\}$ for some $n \in N$, and with subsets of the index set $N$ is used. Thus, for a subset of $\mathcal{G}$ containing goals indexed by some $I \subseteq N$, $\mathcal{G}_I$ is written (see Figure 31).

**Example.** In the BigCorp example: $\mathcal{G} := \{G_0, G_1, ..., G_{10}\}$, where $\mathcal{G}_I$ for $I = \{4, 6, 7, 8\}$ are SoM unit’s goals, i.e. a subset of BigCorp’s goals.

Ideally, the GQM*Strategies goals with related context and assumption elements model the events of interest. This assumes that all relevant parties and organizational units have a shared understanding and agreement regarding what the goal means. Therefore, consensus building among goal owners is very important, and the GQM*Strategies grid derivation process can be iterated until consensus is built.

### 8.2.2 Searching the Event Space: Goal Refinement Patterns and Strategy Types

The GQM*Strategies strategy is fully aligned with the common understanding that a strategy is a guideline for defining goals [48, 122]. However, the formulation and roles of GQM*Strategies strategies are different from goals. Goals are precisely defined and documented with the goal template with eight dimensions (see Sections 2.3 and 6.2.2), while strategies are formulated as statements (see Sections 2.3 and 6.1). Those different levels of formality used to document goals and strategies are keys to a successful search of the event space. Before explaining the search mechanism and how strategies are built and developed, a necessary convention for denoting strategies is defined.

**Strategies** According to the GQM*Strategies meta-model (Section 2.3.1), strategies are used to refine organizational goals; in other words, they address a set of goals and refine them in a set of goals at the next lower level. So, let $\mathcal{G}$ be a set of given organizational goals, $N$ an index set, and $I, J$ sets such that $I \cap J = \emptyset$ and $I, J \subset N$; then the following denotes:

$$\sigma^I(\mathcal{G})_J$$
the strategies defined on a set of organizational goals $\mathcal{G}$ that connect goals $\mathcal{G}_1$ from one level and refine them into the set of goals $\mathcal{G}_1$ at the next lower level.

To simplify, $\sigma(\mathcal{G})^I_J$ is $\sigma(\mathcal{G})^I_J$, if $I = \{i\}$, and as $\sigma(\mathcal{G})^I_J$ if $I = \{i\}$ and $J = \{j\}$. Also, $\sigma(\mathcal{G})^I_J$ is called $i$-strategies, and if there are no other strategies $\sigma(\mathcal{G})^K_J$ such that $K \subseteq J$ it is called a simple strategy or simple $i$-strategy. If there exist strategies $\sigma(\mathcal{G})^I_J$ such that $I \subseteq J$, then $\sigma(\mathcal{G})^I_J$ are the $i$-substrategies of $\sigma(\mathcal{G})^I_J$ (see Figure 31). Two different $i$-strategies $\sigma(\mathcal{G})^I_J$ and $\sigma(\mathcal{G})^I_K$ are disjoint if $J \cap K = \emptyset$.

**Assumption.** It is assumed that each of the $i$-strategies is simple, or there exists a unique decomposition (see Section 8.2.2 below) into its disjoint $i$-substrategies. Later, an input from experts will be required regarding strategies and this assumption will allow experts to provide inputs in a simple and feasible way.

**Example.** BigCorp’s grid (Figure 39 in Appendix 2) shows that top-level goal $G_0$ is refined with strategies into goals $G_1$ and $G_2$ at the next level, i.e. $\sigma(\mathcal{G})^0_{1,2}$. Also, the goal $G_4$ is refined into three goals at the bottom-level, i.e. $\sigma(\mathcal{G})^4_{6,7,8}$. However, the first of the $i$-strategies is not the simple $i$-strategy, because there are two sub strategies $\sigma(\mathcal{G})^0_{1}$ and $\sigma(\mathcal{G})^0_{2}$, while the strategy $\sigma(\mathcal{G})^4_{6,7,8}$ is a simple $i$-strategy.
For the sake of simplicity, it is assumed that a process of defining goals and strategies starts at the top level. Given an event for any high-level unit, modeled with a corresponding goal of that unit, the next step is to identify the supportive events for the next lower level-units, i.e. defining the next lower level units’ goals. This involves identifying a set of potentially equally good goals that support the higher level goal. GQM* Strategies goals and strategies have different roles in the process of developing grids, they both aim at depicting future event(s) but with different roles—selecting a collection of potential events (strategy) vs. selecting a single event (goal). The process of developing grid alternates between these two activities. The mix of these two activities is the main mechanism for searching the event space and defining the next level event(s). Consider the following example. At one level, a goal is defined for a unit, meaning that a future event is specified for the unit. The next step is to define future events for the next lower-level relevant units. By defining strategies for addressing the goal, the options for the next lower-level goals are reduced, i.e. the subspace of the event space is selected. Further, goal owners at the next level will define goal(s) within the selected subspace, by doing that they will specify future events for theirs units. This process of selecting goals from a subspace represents the concept of bounded flexibility that was discussed in Section 6.2.1. This whole process of searching the event space repeats until the entire organizational structure is satisfactorily covered; in other words, it continues until the goals and strategies for all selected organizational units are developed.

Context and assumptions are important in the selection of strategies as they limit the potential event space. The ability to search the event space is one of the strengths of the GQM* Strategies method.

While searching an event space, goals are defined and refined at multiple organizational levels. The process of goal refinement is impacted by strategies. Different strategy types lead to different goal refinement patterns.

During the industry pilots with the GQM* Strategies, several patterns of goals refinement or strategy development were identified. These patterns led to a classification of strategy types. There are two basic types of strategies: AND-Type and OR-Type, while all other strategy types are the result of a combination of these two basic types.

**Required Multiple Strategies: AND-Type Strategies**

When a goal is addressed with several simple i-strategies and when all simple strategies are needed and required for the successful achievement of a goal, then such refinement
patterns can be depicted as the **AND-Type** strategies (Figure 32). The logic of the **AND-Type** strategies can be presented as a conjunction of $i$-strategies, therefore if for the successful achievement of a goal $G_i$ some $i$-strategies $\sigma(G^i_{J_1}), \sigma(G^i_{J_2}), \ldots, \sigma(G^i_{J_k})$ are needed, it is written as:

$$\sigma(G^i_{J_j}) := \sigma(G^i_{J_1}) \land \sigma(G^i_{J_2}) \land \ldots \land \sigma(G^i_{J_k}),$$  \hspace{1cm} (23)$$

where $J := \bigcup_{j=1}^{k} J_j$ expresses the joint effect of $i$-strategies, and $\land$ stands for a conjunction of strategies. The Equation (23) reads: for the successful realization of a goal $G_i$ goals $G_{J_j}$ at the next lower level, refined by strategies $\sigma(G^i_{J_j})$, where $j = \{1, 2, \ldots, k\}$, are all **needed** (i.e. required).

**Example.** For improving process performance (goal $G_2$) experts agreed that a combination of three strategies is needed: adopt agile practices ($\sigma(G^2_{J_4})$) and increase the responsiveness of the customer support process ($\sigma(G^2_{J_1})$, where $j = \{2, 3, \ldots, k\}$) and reduce manufacturing process costs ($\sigma(G^2_{J_5})$). If one of these strategies fails, then it is certain that process performance goal will not be reached.

![Fig 32. Strategy types.](image)

### Alternative Strategies: OR-Type Strategies

Sometimes it is acceptable to define alternative strategies for goals. The benefit of having alternative strategies is risk diversification and an increase of the chances of achieving success. As opposed to **AND-Type** strategies, the alternative strategies are not all **needed**, i.e. they represent redundant ways to achieve the same goals; and thus, they reduce risks on the expense of increasing resources for strategies implementation.
Formally, the logic of the OR-Type strategies can be presented as a disjunction of \( i \)-strategies; therefore, if for the achievement of a goal \( G_i \) some alternative \( i \)-strategies \( \sigma(\mathcal{G})_{J_1}^i, \sigma(\mathcal{G})_{J_2}^i, ..., \sigma(\mathcal{G})_{J_k}^i \) are given, it is written as:

\[
\sigma(\mathcal{G})_{J}^i := \sigma(\mathcal{G})_{J_1}^i \lor \sigma(\mathcal{G})_{J_2}^i \lor ... \lor \sigma(\mathcal{G})_{J_k}^i,
\]

where \( J := \bigcup_{j=1}^k J_j \) stands for the joint effect of \( i \)-strategies, and \( \lor \) stands for the disjunction of strategies. The Equation (24) reads that only one set of goals \( \mathcal{G}_{J_j} \) refined by alternative strategy \( \sigma(\mathcal{G})_{J_j}^i \), for some \( j \in \{1, 2, ..., k\} \), is sufficient for the successful realization of a goal \( G_i \).

In practice, the GQM+Strategies grid specifies the logic of a goal refinement pattern with context and assumption elements. However, in the case of large grids, analyzing types of strategies takes a significant mental effort; therefore, the use of a visual notation for marking alternative strategies is suggested while a GQM+Strategies grid is developed, as shown in Figure 32 for the OR-Type strategies.

As a result of a single strategy, a set of goals can be derived from the upper-level goals (Section 2.3); therefore, it is important to discuss goal refinement patterns with a single strategy.

**Goal Refinement Patterns with a Single Strategy**

Figure 33 depicts all possible cases of goal refinement using a single strategy. The simplest form is a *singleton simple \( i \)-strategy* (33.a), where a strategy refines a single goal into another goal at the next level.

Another case is a situation that occurs when a single simple strategy addresses one goal and generates a collection of next-level goals (Figure 33.b). The presumed fact is that all lower-level goals are needed for the successful achievement of the upper-level goal, and it is hard or impossible to distinguish their individual contributions. The collective impact of the generated goals on the addressed goal is synergistic, meaning that the joint impact is larger than a simple sum of individual impacts.

**Note.** When a simple strategy generates a collection of goals, all those goals are needed and required. Alternative goals are generated by alternative strategies (see Section 8.2.2).

Figure 33.c.1 illustrates the most general case for a single strategy refinement, as defined in Section 8.2.2, when a collection of the lower-level goals \( \mathcal{G}_{J} \), refined by a
Fig 33. Special cases of strategy types. (a): Singleton basic i-strategy. (b): A single strategy $\sigma(G)^J$, where $J = \{2, 3, ..., n\}$, leads to a set of goals $\mathcal{G}_J$ at the next level. (c.1): A single strategy $\sigma(G)^I$ addresses a set of goals $\mathcal{G}_I$ and refines them to a set of goals $\mathcal{G}_J$ at the next level. This case is equivalent to the case (c.2) of having $k$ times i-strategy for goals $G_i, G_j \in \mathcal{G}_J$. A single strategy $\sigma(G)^J$, has impact on a collection of the upper-level goals $\mathcal{G}_I$. However, the upper-level goals are independent (otherwise, they would not be at the same level), and it is hard to perceive that failing to achieve some upper-level goals would imply the failure of the remaining upper-level goals from the same level. Therefore, this situation is transformed into an equivalent case, where each upper-level goal is addressed with i-strategies that result with same set of lower-level goals (Figure 33.c.2).

Derived Form of Strategies

It is assumed that the achievement of goal $G_i$ is the result of composing all its successful $i$-strategies. Each of the $i$-strategies in the composition is either simple or derived via the disjunction or conjunction of its disjoint $i$-substrategies. For a goal $G_i$, the composition represents what is called an $i$-superstrategy, denoted by $\Sigma^i$ (capital Greek letter sigma), to differentiate it from $i$-strategies denoted by small Greek letter sigma.
**Example.** For BigCorp’s goal of improving process performance \( (G_2) \) the \( i \)-superstrategy \( \Sigma^2 \) is comprised of three different strategies: adopt agile practices, increase responsiveness of customer support process, and reduce manufacturing process costs. All these strategies are needed for the successful achievement of goal \( G_2 \).

### 8.2.3 Strategy Effectiveness

Due to the inherent uncertainty that exists in the environment, it is obvious that some strategies that comprise an \( i \)-superstrategy will be more or less successful than other strategies. For a simple \( i \)-strategy, an expert opinion can be elicited regarding strategy effectiveness as a way of measuring the **strategy effectiveness**. Strategy effectiveness is the degree to which a strategy delivers the planned results. Expert feedback is only required for simple strategies; the input will be used to calculate the effectiveness of the superstrategies (as explained in the following sections).

For a simple \( i \)-strategy \( \sigma(\mathcal{G})^i_j \), an expert assessment regarding its effectiveness can be probed with the following questions:

- \( Q_i \): What is the likelihood of having a **successful strategy** \( \sigma(\mathcal{G})^i_j \) if all derived **goals** from the strategy, \( \mathcal{G}_j \), are already **achieved**?

A successful strategy is a strategy that delivers the results (effects) that are expected by the experts. Those results in combination with other results from different strategies should lead to the successful realization of a goal for which the strategies were defined.

Note that the focus of the question \( Q_i \) is on a strategy and on the effects of the strategy, not on the prediction of a goal’s achievement.

Analyzing the grid derivation process and how it relates to the concepts of events and event space helped to identify the main question during the grid derivation process; therefore, the experts have a chance to provide the best possible answer to this question. In later sections, strategy effectiveness will be operationalized to measure and define the question as an instrument for collecting expert assessments regarding strategies’ effectiveness.
8.3 Solution Part II: The Causal Reasoning over the GQM+Strategies Grids

In this section, the second subproblem is addressed through a mathematically sound mechanism used to calculate the probability distribution values from the pieces of information that are provided by experts. This will require a mathematically formal notation of the GQM+Strategies concepts. Before doing that, the needed mathematical definitions and notations are given in the following section.

8.3.1 Basic Mathematical Definitions and Notation

A representation of the causal models with graphs and the relationship between probability distribution and causal graphs (i.e., Bayesian network) is the main background theory for the second subproblem. A book on Causality by Pearl [166] was the primary source. All definitions and notation are adopted from the book, if not noted otherwise. In what follows, the mathematical notation and terminology of the concepts of directed acyclic graphs, random variables, joint probability distribution, and causal graphs will be defined.

Directed Acyclic Graphs

The structure of the GQM+ Strategies grid can be formally defined as a specific type of graphs, i.e. a directed acyclic graph (DAG).

A directed graph, or simply a digraph, is an ordered pair of a set of nodes and a set of arcs (a set of ordered pairs of the node set). Finite digraphs without loops, i.e. without arcs of type \((x, x)\), where \(x\) belongs to the node set of a graph, will be dealt with.

Let \(D = (V, A)\), be a digraph, then \(A \subseteq V \times V\). A sequence of nodes from \(V\), \(P := (v_0, v_1, ..., v_k)\) is a directed path, or simply a dipath, in a graph if \((v_i, v_{i+1}) \in A\) for all \(0 \leq i < k\). If \((u, ..., v)\) is a dipath, also called a uv – path, and it is said that \(v\) can be reached from \(u\). Also, if \((u, v) \in A\), \(u\) is called a parent of \(v\), and \(v\) is called a child of \(u\). If \(v\) can be reached from \(u\), then \(u\) is an ancestor of \(v\), and \(v\) is a descendant of \(u\). The digraph \(D\) is acyclic (DAG), if there are no two nodes \(u, v \in V\) where \(v\) can be reached from \(u\), and \(u\) can be reached from \(v\).
Random Variables

The GQM* Strategies goals and strategies can be associated with specific events that occur within an organization, and some of the events or sets of events affect the occurrences of other events. Therefore, the concept of random variables and sets of random variables is used to represent the organizational events.

Next, consider a set of random variables \( X = \{X_0, X_1, \ldots, X_n\} \), \( n \in \mathbb{N} \), in some event space (i.e. some body of knowledge\(^{37}\)) equipped with some probability distribution \( P \). Set \( N := \{0, 1, 2, \ldots, n\} \), where \( i \in N \), \( S_i \) denotes the domain of \( X_i \); if \( S_i \) is finite for \( x_i \in S_i \), the situation \( "X_i \) is realized as \( x_i" \) (i.e. \( "X_i \) gets value \( x_i" \)) is an **elementary event** of \( X_i \), and it is expressed by \( X_i = x_i \). It is assumed that the random variables from \( X \) have finite domains, hence there is an "one to one" correspondence between the **elementary events** of a random variable and the elements of its domain set, which allows to work with the entire sets of random variables and sets of events instead of working with individual variables.

The progress and success of an organization can be uniquely depicted by assessing the achievement of organizational goals in the GQM* Strategies grid. For a single goal possible situations can be **achieved or not achieved**, and those are the examples of elementary events.

Through this section the subsets of \( X \), i.e. with subsets of the index set \( N \) are used. For \( I \subseteq N \), set \( X_I := \{X_i \mid i \in I\} \), and let \( x_I := \{x_i \mid i \in I\} \) be such that \( x_i \in S_i \) for all \( i \in I \), i.e. \( x_I \in S_I := \bigcap_{i \in I} S_i \). Also, \( X_I = x_I \) is used for the set of events \( \{X_i = x_i \mid i \in I\} \). Moreover, if \( I \) is a singleton, i.e. \( I = \{i\} \) for some \( i \in N \), to simplify, \( X_i, x_i \) is used instead of \( X_{\{i\}}, x_{\{i\}} \). Also, \( P(x_{I_1} = x_{I_1}, x_{I_2} = x_{I_2}, \ldots, x_{I_k} = x_{I_k}) \) is denoted by \( P(x_{I_1}, x_{I_2}, \ldots, x_{I_k}) \), and represents the probability of occurring \( X_{I_1} = x_{I_1}, X_{I_2} = x_{I_2}, \ldots, X_{I_k} = x_{I_k} \). In other words, \( X_{I_1} = x_{I_1}, X_{I_2} = x_{I_2}, \ldots, X_{I_k} = x_{I_k} \) represents \( k \) sets of events and \( P(x_{I_1}, x_{I_2}, \ldots, x_{I_k}) \) is the probability or joint probability of the events occurring together.

For \( I, J \subseteq N \), the probability of \( X_I = x_I \) occurring if \( X_J = x_J \) occurred is called a **conditional probability**, and is written as \( P(X_I = x_I|X_J = x_J) \), or simply \( P(x_I|x_J) \). Formally:

\[
P(x_I|x_J) := \frac{P(x_I, x_J)}{P(x_J)}, \tag{25}
\]

where \( P(x_I, x_J) \) denotes the probability of both \( X_I = x_I \) and \( X_J = x_J \) occurring.

---

\(^{37}\)The events are modeling the real world, and therefore they represent knowledge about that real world.
Let $I, J, K \subseteq N$, then $X_I$ and $X_J$ are **conditionally independent** given $X_K$, if:

$$P(x_I \mid x_J, x_K) = P(x_I \mid x_K)$$

for all $x_I, x_J$, and $x_K$ such that $P(x_J, x_K) > 0$. If $K = \emptyset$, then $X_I$ and $X_J$ are **marginally independent**.

If one set of random variable is *marginally independent* from another set of random variables, then the occurrences of the events defined by the first set of random variables are not impacted or affected by the events of the second set of variables. In a GQM* Strategies grid this will always occur when some goals are not affected by some other goals and strategies in the grid.

*Conditional independence* introduces a third set of random variables, which represents a situation when the occurrences of the events of the first set of random variables are not affected by occurrences of events of the second set when second-set events are observed together with the events of the third set. In other words, there might be some dependencies between the first and second set, but the third set is canceling those dependencies; therefore, the third set defines the condition that renders the first set *independent* from the second set of events. For example, in a grid, some goals and strategies can affect other goals in a negative way (i.e. define a set of contradicting goals). In this situation, one possibility is to define a new set of strategies and goals that will neutralize contradicting strategies and goals. That new set of goals and strategies defines a condition that renders goals independent from the contradicting goals.

If the event space for random variable $X$ is $K$, one may be more precise and write $P(X = x \mid x \in K)$ instead of $P(X)$. Throughout, assume $P(X) = P(X \mid K)$, and $K$ will be omitted if the context allows. The conditional probability is one of the fundamental principles used by Bayesian philosophers. Further, Bayesianists advocate that any probabilistic reasoning about real-world events is conditionally dependent on an existing body of knowledge (also referred to as the context). For example, if $A$ is an event of interest and $P(A)$ is the probability of that event, a Bayesian philosopher would claim that a fundamental fact—a body of knowledge—was missed, and the probability should be expressed as $P(A \mid K)$, where $K$ is the body of knowledge or, in other words, the probability of an event $A$ with a given body of knowledge $K$. For simplicity, writing conditional dependencies from a body of knowledge is omitted.
The Joint Probability and the Causal Graph

Desirable, e.g., organizational, events can be defined using a set of random variables $X_N$, then the probability of event $x_N$ or joint probability $P(x_N)$ can be calculated using the following identity:

$$P(x_N) = \prod_{i=0}^{n} P(x_i | x_{i+1}, x_{i+2}, ..., x_n), \quad (26)$$

where $x_N = \{x_0, x_1, ..., x_n\} \in S_N$, known as the chain rule of probability calculus, can be easily derived from the definition of conditional probability defined with Equation (25).

Using the chain rule for calculating the joint probability requires knowledge about dependencies among various sets of events, and often this knowledge is hard to obtain. Fortunately, the nature of the problem offers a simpler way to calculate the joint probability using a smaller set of variables than with the chain rule. The criteria for selecting that subset of variables is determined by Markovian parents, as described in the following.

If for $i \in N$, and set $m(i) \subseteq [i + 1, i + 2, ..., n]$, when the following holds:

$$P(x_i | x_{m(i)}) = P(x_i | x_{i+1}, x_{i+2}, ..., x_n) \text{ for all } x_i, x_{i+1}, ..., x_n, \quad (27)$$

where no proper subset of $m(i)$ satisfies Equation (27), the elements of $X_{m(i)}$ are called Markovian parents of $x_i$.

The consequence of the Equation (27) is relevant because for the GQM* Strategies goal is possible to identify a set of goals that are impacting the goal, and at the same time, the identified set of goals isolates the goal from the impacts of the remaining goals and strategies in the grid.

This is significant because of the tree-like structure of the GQM* Strategies grid allows the identification of Markovian parents in a trivial and simple way. This is going to be a starting point for defining the organizational causal graphs.

If the function $m : N \rightarrow P(N)$ is such that $m(i) \subseteq [i + 1, i + 2, ..., n]$ and $X_{m(i)}$ are Markovian parents of $x_i$, then DAG $D = (X, A)$ defined by $A := \{(X_i, X_j) \mid X_i \in X_{m(j)}\}$ is called a causal graph or Bayesian network. Note that $D$ has the set $X$ as the node set and there is an arc from $X_i$ to $X_j$ if $X_i$ is a Markovian parent of $X_j$. Now, Equation (26) reads as:

$$P(x_N) = \prod_{i=0}^{n} P(x_i | x_{m(i)}), \quad (28)$$

\[38\] It is known that this function is unique whenever $P(x_N)$ is strictly positive for all $x_N \in S_N$. 

207
and it is said that $P$ admits factorization relative to DAG $D$. In general, many probability distributions provided for a given set of random variables can share (have) a same Bayesian network. If a probability distribution $P$ admits the factorization relative to a DAG $D$, then the DAG $D$ and the probability distribution $P$ are compatible\(^{39}\).

The necessary and sufficient conditions for the compatibility of a joint probability distribution $P$ and a DAG $D$ are given with a theorem in [166, p. 120]. According to the theorem, a probability distribution $P$ over $X$ and DAG $D = (X, A)$ are compatible if and only if every variable from $X$ is rendered conditionally independent, by its parents, from all other variables that are not its descendants. The compatibility condition will allow specific input probabilities (i.e. probabilities defined by expert) through the analysis of goals and strategies that are connected to the goal, instead of analyzing the entire grid at once.

If $D$ and $P$ are compatible, as a major consequence is that $P$ can be assessed in factored form (28). This is important because assessing (specifying) probabilities required for the factored form is much easier and does not require specifying probabilities for each combination of the sets of dependent variables, as it would require for the chain rule.

Furthermore, the assessment of $P$ via Equation (28) enables the calculation:

$$P(x_I) = \sum_{x_K} P(x_I, x_K),$$

(29)

for $I \subseteq N$ and $K := N \setminus I$, and also

$$P(x_I \mid x_J) = \frac{\sum_{x_K} P(x_I, x_J, x_K)}{\sum_{x_L} P(x_J, x_L)}$$

(30)

for $I, J \subseteq N$, where $K := N \setminus (I \cup J)$ and $L := N \setminus J$. Equation (30) represents an inference mechanism over a Bayesian network, which stands at the core of the causal reasoning. Causal reasoning will be used to run different queries over the GQM\(^{+}\)Strategies grid.

8.3.2 The GQM\(^{+}\)Strategies Formal Notation

In this section, a formal notation of the remaining GQM\(^{+}\)Strategies concepts is introduced. The formal notation will allow the use of a mathematical device in a formal and rigorous way.

\(^{39}\)It is common to say that a DAG represents distribution, or that distribution is Markov relative to a DAG.
Organizational Units, Goals, and Contextual Variables

An organization is immersed in an environment. The external environment can be differentiated from the internal environment. The most natural way to define the external environment is as circumstances beyond corporate boundaries (e.g., the market, technology), while the internal environment is a characterization of an organization itself. This differentiation of environments is important, because the internal environment is often neglected, and more attention is given to the external environment. The necessary information about environments is provided by experts.

The term expert is used to refer to a knowledgeable representative from an organization. Their knowledge and understanding of the organizational business and operations (i.e. environment), opinions, and estimates are important for the successful application of the approach. Enterprises usually have a hierarchical organizational forms. Therefore, a hierarchy of organizational units will be used as a representation of an enterprise. An organizational unit can be a company within a larger corporation; the company has its departments as sub-units.

Example. BigCorp’s organizational structure is plotted in Appendix 2. It contains several levels of units hierarchically organized, from management board ($U_0$) representing strategic or top-level, to the mid-level four business units: marketing & PR ($U_1$), SoM ($U_2$), MHP ($U_3$), and customer support ($U_4$). Each business unit contains several sub-units, like R&D ($U_6$) is a sub-unit of the SoM business unit.

The activities of the units are mainstreamed with organizational goals or just goals. Goals can be defined at different organizational levels and within different units. Let $\mathcal{G} = \{G_i \mid i = 0, 1, \ldots, k\}, k \in \mathbb{N}$ be a set of organizational goals. Since each goal is related to its unit, this is comprised by a function $u : \mathcal{G} \rightarrow \mathcal{U}_G$, where $\mathcal{U}_G$ is the set of organizational units of goals from $\mathcal{G}$. The organizational unit of a goal $G \in \mathcal{G}$ is referred as $u(G) \in \mathcal{U}_G$. Although $u$ is surjective by definition, it is not necessarily injective. Namely, there may be more than one goal assigned to a same unit.

The environment (external and internal) represents circumstances that can be depicted with contextual variables. A particular instance of internal and external contextual variables can be used to characterize a position of a unit with respect to its

---

40A function $f$ from a set $X$ to a set $Y$ is surjective (or onto), or a surjection, if every element $y$ in $Y$ has a corresponding element $x$ in $X$ so that $f(x) = y$. Multiple elements of $X$ might be turned into the same element of $Y$ by applying $f$. 

209
goals. Let denote the set of all relevant contextual variables for goals from $G$ by $\forall G$, and for $v \in \forall G$ where $D_v$ denotes its domain. To each goal $G \in G$, which is assigned a subset of $\forall G$ whose elements are all relevant contextual variables for $G$. It is assumed that these variables change their values as time passes\textsuperscript{41}, and their values provide all relevant information regarding $u(G)$ with respect to $G$, i.e. information regarding the progress of a unit toward its goals is gathered. This can be described as function $c : G \rightarrow \mathcal{P}(\forall G)$, where $\mathcal{P}(\forall G)$ denotes the power set of $\forall G$. In other words, function $c(\cdot)$ returns a set of relevant contextual variables for a goal $G$, where $G \in G$. If the function $c(\cdot)$ is injective (which is often the case in a real-world situation), and if $V = c(G)$ for some $G \in G$, one can verify that $u(c^{-1}(V)) = u(G)$ and observe that $V$ is thus directly linked to unit’s goal $u(G)$. Note that the GQM+Strategies goal template explicitly specifies an organizational unit as the scope (context) dimension of a goal, which makes it possible to identify the unit of a goal, so to which unit a goal belongs can be identified, and the goal itself is sufficient to set up a relevance criteria for contextual variables.

**Position and State**

Let $G$ be a goal, set $U := u(G)$ and $V := c(G)$. The elements of $\otimes_v D_v$, where $\otimes$ denotes the Cartesian product, will be called the positions of $U$ with respect to $G$.

Determining an exact position of an organizational unit, i.e. determining an appropriate element in $\otimes_v D_v$ in order to describe $U$ with respect to $G$ is a challenging task due to the nature of contextual variables. For example, an assessment of a goal to increase the market awareness about some product is more difficult to express with exact numbers than with categories of values (e.g. low, medium, high). Therefore, the partitions of domains are introduced as a more convenient concept, where partitions of domains should be designed and chosen in such way that assessing current states becomes easy.

To each variable $v \in V$ a partition of $D_v$ denoted $S_v$ is assigned, i.e., a collection of disjoint subsets of $D_v$, such that their union equals $D_v$. A product of subsets $\otimes_v S_v$, such that $s_v \in S_v$ for all $v \in V$ is a state of $U$ with respect to $G$. Hence, a state can be viewed as a collection of subsets of values of contextual variables. To simplify the notation, $\mathcal{S}_G := \otimes_v S_v$ is used for the set of all states of $U$ with respect to $G$, and small letters, e.g. $s, s_1, \ldots$, are used for states. Also, $\mathcal{S}_G$ is called the state space of $U$ with respect to $G$, and for a state $s$ is written $s \in \mathcal{S}_G$.

\textsuperscript{41}Note that contextual variables that are not known for sure are represented with predictions or assumptions.
Time Component of States and Goals

The entire concept of goals is not only organizational-unit dependent, it is also time dependent, where the GQM+Strategies goals can be viewed as guiding markers or milestones of a larger process. It is the process of positioning organizational units into certain states in certain moments of time. Therefore, a state can be a desired future situation—future state, characterization of a current situation—current state, or a previous situation—past state of some organizational unit. For example, a future state can be positioning a corporation in the market as a recognized market-leader for a certain type of services. Yet another example would be situating the R&D department in a position to deliver more innovative concepts in shorter time periods. Also, a unit can have several goals, and this is the situation when an organization or its units require positioning into several states simultaneously. Those states are referred to as the concurrent states of a unit.

For each goal, its unit is formally described via function $u(\cdot)$, and its context via function $c(\cdot)$. Next, the time aspects of a goal is defined since its achievement is always related to a time horizon or timeframe. Note that the GQM+Strategies goal template specifies timeframe as one of the dimension of a goal. Let $T_G$ be the set of all (relevant) time moments for goals from $G$. Therefore $t : G \rightarrow T_G$ can be defined, and for a goal $G \in G$, the $t(G)$ is called the time horizon of $G$.

The importance of time horizons is unavoidable in the GQM+Strategies grid derivation process. For two goals $G_i$ and $G_j$ such that $G_j$ refines $G_i$ it can be assumed that $t(G_j) \leq t(G_i)$. Otherwise, accomplishing $G_j$ at $t(G_j)$ will not have any influence on $G_i$, and consequently, refining the goal $G_i$ with $G_j$ would not make any sense.

Events and Random Variables

In the GQM+Strategies grid, each goal is defined with intention and belief that the goal is or will be valuable for the organizational unit at a specific point of time. In other words, the architects of a grid (i.e. experts) have certain expectations regarding goals and results that goals will deliver. Formally, this can be represented using random variables.

Let $G_i \in G$, set $U := u(G_i)$, $V := c(G_i)$, $T := t(G_i)$. The positioning of $U$ into a state $s \in S_G$, at the moment $T$ is an event. For each event, a belief\(^{42}\) is accepted about its

---

\(^{42}\)Experts opinion.
occurrence. It is possible to express the beliefs by

\[ P_{G_i} : \mathbb{S}_{G_i} \rightarrow [0,1], \]

where the \( P_{G_i}(s) \) is the belief for a unit \( U \) to be in a state \( s \) at the moment \( T \). Since \( U \) can be in exactly one of the states \( \mathbb{S}_{G_i} \) at \( T \) [with respect to a goal \( G_i \)], without loss of generality, it is assumed that \( P_{G_i} \) is a probability distribution, i.e. \( \sum_{s \in \mathbb{S}_{G_i}} P_{G_i}(s) = 1 \).

Next, for a goal \( G_i \in \mathcal{G} \) a random variable \( X_i \) can be assigned, where \( i \) is index of a goal in \( \mathcal{G} \) that takes values from \( \mathbb{S}_{G_i} \) according to \( P_{G_i} \). Hence, a random variable \( X_i \) models events (e.g., outcomes, realizations) that occur according to \( P_{G_i} \). In other words, a random variable \( X_i \) models the event of positioning a unit \( U \) into states with respect to the goal \( G \).

The event of interest is the successful positioning into a desired state, which (if not stated otherwise) indicates the achievement of a goal. Therefore, for a goal \( G_i \) possible outcomes/states are "\( G_i \) is achieved" denoted \( \{X_i = x^*_i\} \), and "\( G_i \) is not achieved" denoted \( \{X_i = \bar{x}_i\} \); and, they represent the elementary events of the random variable \( X_i \). For the set of all random variables of goals from \( \mathcal{G} \) a set \( \mathcal{X}_G \), i.e. \( \mathcal{X}_G := \{X_i \mid G_i \in \mathcal{G}\} \) will be used.

Next, it is necessary to understand how the probabilities \( P_{G_i}(X_i = x^*_i) \) and \( P_{G_i}(X_i = \bar{x}_i) \) for \( X_i \in \mathcal{X}_G \) can be determined, since it will provide all the necessary information about goals dependencies.

**Defining Strategies and "Leads to" Relation**

The GQM\( ^* \) Strategies grid is more than a simple collection of organizational goals and strategies, it is a scheme where certain goals and strategies are put together with an expectation that they can "work" together in a synchronous way and deliver results. Therefore, the entire GQM\( ^* \) Strategies grid can be analyzed by selecting all pairs of goals in a grid and determining the dependency relationship for each pair. Of course, the dependency relationship is the strongest among goals that are connected with a single strategy, but it can exist among other goals as well.

The dependencies among goals in a grid correspond to the dependencies among events that those goals characterize. An event was defined as a situation when one organizational unit is positioned into some state at some moment in time. Events are, in general, influenced by each other; one event can contribute or even prevent another event. Generally, these influences are modeled as dependencies among random variables,
where, the all relevant information regarding goal dependencies are determined by the GQM* Strategies grid development process.

Here, a relationship driven by strategies is introduced to model influences between sets of random variables. As before, a set of goals $\mathcal{G} = \{G_i | i \in N\}$, where $N = \{0, 1, 2, ..., n\}$ is considered. Let $\mathcal{X}_G$ be the set of corresponding random variables. For a subset $I$ of $N$ the $\mathcal{G}_I := \{G_i | i \in I\}$ and $\mathcal{X}_I := \{X_i | i \in I\}$ are used. If $I = \{i\}$ then it is simply written as $X_i$ instead of $\mathcal{X}_i$.

If $G_i$ is a goal, and there exists a set of goals $J$ such that $\sigma(G)_i^j$ is an i-strategy, this is associated with an ordered pair of a random variable and a set of random variables, denoted as:

$$\sigma(G)_i^j : \sim (X_i, \mathcal{X}_j),$$

where the $L : \sim R$ denotes that $L$ is evaluated with $R$. The relationship associated with $\sigma(G)_i^j$ is simply a subset of $\mathcal{X}_G \times \mathcal{P}(\mathcal{X}_G)$, where $\mathcal{P}$ is the power set.

For two random variables, $X_i, X_j \in \mathcal{X}_G$, such that there exists a set of goals $J$ containing goal $j$ with $\sigma(G)_i^j$ being an i-strategy, can be written $X_j \overset{\sigma(G)_i^j}{\sim} X_i$, and said $X_j$ leads to $X_i$, which evaluates the probability that $G_i$ is achieved. Note that $\overset{\sigma(G)_i^j}{\sim}$ is a binary relation on $\mathcal{X}_G$.

The significance of the leads to relation is that as a binary relation it can be used to model different types of dependencies between two goals, such as contradicting goals.

**The Universe of Discourse and GQM*Strategies Grid Structure**

The objective of this entire section is to introduce a formal notation of the GQM* Strategies concepts. First, to formalize the basic concepts of the GQM* Strategies, it is important to discuss how they relate. Now, the big picture will be formed, i.e. define the GQM* Strategies grid as a whole. Before defining formally the grid structure, an abstraction of the real-world, for which all the grids are derived, must be defined. The abstraction of the real-world is referred to as the universe of discourse.

The universe of discourse represents a space that consists of all potential situations and their dependencies. The situations, in which an organization can be positioned, are defined by all possible instances of contextual variables. The complexity of modern enterprises suggests that various parts of an organization form their own sub-universes of events in which they are immersed.
Let \( G \) be a set of goals indexed by \( N \), \( \sigma(G) \) are all strategies defined over \( G \), and \( X_G \) represents all random variables associated with goals \( G \). The pair is defined:

\[
\mathcal{I}_G := (X_G, \sigma(G))
\]

as the universe of discourse. In other words, the universe of discourse represents all possible strategies and goals for an organization and its units and associated potential events that evaluate strategy effectiveness and goal achievement.

If \( I \subset N \), and \( \sigma(G_I) := \{(X_j, X_K) \mid j \in I \setminus K, \ K \subseteq I\} \), then \( \mathcal{I}_I := (X_I, \sigma(G_I)) \) is a subuniverse of \( \mathcal{I}_G \) generated by \( G_I \).

According to the GQM\(^+\)Strategies meta-model (Section 2.3.1), the achievement of organizational goals and strategies is assessed through the GQM measurement scheme. The role of the measurement scheme is to provide evidence, i.e. data, of achieving organizational goals. For the purpose of completing the formal notation of the GQM\(^+\)Strategies concepts, a GQM based measurement scheme is defined as a hierarchical collection of measurement goals, questions, metric definitions, and interpretation models, and it is denoted by \( M_G \). The GQM based measurement schemes are attached to the organizational goals from \( G \).

Finally, all elements of the GQM\(^+\)Strategies grid have been formally defined, and the tuple:

\[
\Xi_G := (X_G, \sigma(G), M_G) \equiv (\mathcal{I}_G, M_G)
\]

represents the GQM\(^+\)Strategies grid. Grid \( \Xi_I = (\mathcal{I}_I, M_I) \) is also referred to as a sub-grid of a grid \( \Xi_G \), noted as \( \Xi_I \subseteq \Xi_G \), if \( I \subset N \), \( M_I \subseteq M_G \) and \( \mathcal{I}_I \) is a sub-universe of \( \mathcal{I}_G \).

The impact of the environment on a strategy definition and selection is modeled by all relevant dependencies among events, as defined by the universe of discourse. Considering an GQM\(^+\)Strategies grid and its formal representation \( \Xi_G \), the context and assumption elements and their influences (relationships) on goals and strategies are modeled with the universe of discourse \( \mathcal{I}_G \) as they define current and future events of interest. A successful business has to be able to search the universe of discourse to identify desired future events for the whole organization and navigate through such a subset of events by achieving its objectives.

### 8.3.3 Causal Reasoning based on the GQM\(^+\)Strategies Grid

Any probabilistic interpretation of a given grid requires defining a probability distribution over that grid. Let \( G \) be the set of goals indexed by \( N \), and its grid \( \Xi_G = (X_G, \sigma(G), M_G) \).
In Section 8.2.2, how the strategies affect different goal refinement patterns was discussed, and according to Section 8.3.2, a strategy \( \sigma(G)_i^j \) captures influences between \( X_j \) and \( X_i \). On the other hand, the DAG \((X_G, \sim)\) captures causal relations among goals in \( G \), i.e. among random variables from \( X_G \). In this section, a probability distribution will be defined based on expert knowledge about strategies and their effectiveness. The section is concluded with the observation that such a distribution is compatible with \((X_G, \sim)\), which represents an organizational causal model.

### Conditional Dependencies Among Goals and Strategies

In order to quantify strategy effectiveness, the conditional probability can be used. Let \( G_i \in G, J \subset N \) such that \( \sigma(G)_i^j \) is an i-strategy. There are only two outcomes for a strategy: it can either be successful, denoted as \( \sigma(G)_i^j = \tilde{\sigma}_i^j \), or fail, denoted as \( \sigma(G)_i^j = \bar{\sigma}_i^j \). Hence, \( \sigma(G)_i^j \) can be viewed as a random variable\(^{43}\), and the measure of the effectiveness of \( \sigma(G)_i^j \) is defined as \( P(\sigma(G)_i^j = \tilde{\sigma}_i^j) \), i.e. the probability that the goal \( G_i \) is achieved if strategy \( \sigma(G)_i^j \) was successful.

In Section 8.2.3, it is assumed that experts can provide information about the expectancy of the successful realization of a simple i-strategy, assuming that they know the information about goals achievement for all \( j \in J \). Basically, it means that experts provide:

\[
P(\tilde{\sigma}_i^j | x_j) \text{ or } P(\bar{\sigma}_i^j | x_j),
\]

for realization \( x_j \) of variables from \( X_j \), i.e. conditional probabilities for \( \sigma(G)_i^j \) being successful \( (\tilde{\sigma}_i^j) \) or not being successful \( (\bar{\sigma}_i^j) \) if, for example, all derived goals from strategy \( \sigma(G)_i^j \) are achieved \( (x_J = x_J^*) \).

In what follows, causal independence will be used, see e.g. [114] for details, also known as noisy functional dependence, to specify interactions among causal variables, and give compact formula for conditional probabilities for all i-strategies.

Let an i-strategy be either basic, or of the AND-type or the OR-type. If a strategy \( \sigma(G)_i^j \) is of the AND-type, then there exist some i-strategies \( \sigma(G)_i^j_1, \sigma(G)_i^j_2, \ldots, \sigma(G)_i^j_k \), where \( J = \cup_{m=1}^k J_m \). Moreover, they are all needed in order to make \( \sigma(G)_i^j \) successful. For a realization \( x_J \) of variables from \( X_J \), the conditional probability for \( \tilde{\sigma}_i^j \) is defined as:

\(^{43}\)random variables over the same event space as for goals
\[ P(\bar{\sigma}^i_j \mid x_J) := \prod_{j=1}^{k} P(\bar{\sigma}^i_{J_j} \mid x_{J_j}). \] (32)

Trivially, \( P(\bar{\sigma}^i_j \mid x_J) = 1 - P(\sigma^i_j \mid x_J) \).

If a strategy \( \sigma(\hat{G}^i_j) \) is of the OR-type, then there exist \( i \)-strategies \( \sigma(\hat{G}^i_{J_1}), \sigma(\hat{G}^i_{J_2}), \ldots, \sigma(\hat{G}^i_{J_k}) \) where \( J = \bigcup_{m=1}^{k} J_m \), such that the success of any of them would cause the success of \( \sigma(\hat{G}^i_j) \). In other words, \( \sigma(\hat{G}^i_j) \) fails if all strategies \( \sigma(\hat{G}^i_{J_1}), \sigma(\hat{G}^i_{J_2}), \ldots, \sigma(\hat{G}^i_{J_k}) \) fail. Thus, for a realization \( x_J \) of variables from \( X_J \), the following holds:

\[ P(\bar{\sigma}^i_j \mid x_J) := \prod_{m=1}^{k} P(\bar{\sigma}^i_{J_m} \mid x_{J_m}), \] (33)

i.e. it is 1 if all \( \sigma(\hat{G}^i_{J_m}) \) fail, where \( P(\bar{\sigma}^i_j \mid x_J) = 1 - P(\sigma^i_j \mid x_J) \) for OR-type \( i \)-strategies.

Observe next that Equations (32) and (33) with the information provided by experts, in Equation (31), are sufficient for computing conditional probabilities for the \( i \)-superstrategy \( \Sigma^i \). Since its success causes the achievement of the goal \( G_i \), and since in case of its failure \( G_i \) cannot be achieved the following holds:

\[ P(x^*_i \mid x_{J(i)}) := P(\bar{\Sigma}^i \mid x_{J(i)}) \text{ and } P(\bar{x}_i \mid x_{J(i)}) := P(\bar{\Sigma}^i \mid x_{J(i)}), \] (34)

where index \( J(i) \) denotes the complete set of goals refined from \( i \). In this way, a recursive recipe for computing \( P(x_i \mid x_J) \) is acquired, when provided with expert opinions.

**Unobserved simple strategy** In practice, the experts might have limited access to information; therefore, might not be able to observe all possible strategies. This phenomenon can be modeled using an unobserved simple strategy for each goal. Such OR-type strategies should have small probabilities, and they represent error variables. For example, when due to new circumstances the upper-level goal is achieved despite the failure of all strategies defined by experts for that goal, such situation is not unlikely, but the probability of that occurring is very small, and it is associated with the effects of an unobserved simple strategy. However, for the dissertation ideal experts who are capable of observing all possible strategies are assumed.

---

\(^{44}\)Simon defines that as bounded rationality (Section 2.5.1).
Causal Graph and Joint Probability Distribution Function

Using the notation introduced in previous sections, a causal graph can be constructed based on the GQM+ Strategies grid (Figure 34). The causal graph and inputs provided by experts will be sufficient for defining the probability distribution \( P \) over a grid. Beside the nodes that represent organizational goals (white nodes), there is also a need for additional nodes that represent simple strategies (black nodes). Introducing those strategy nodes allows representation of the compound effects of strategies on goals.

Finally, conditional probabilities defined in Equation (34) can be used to construct probability distribution \( P \) over \( X_G \) and promote the DAG \((X_G, \rightarrow)\) into a Bayesian network. For a goal \( G_i \), let \( J(i) \) denote the set of all goals refined from \( i \). For all \( x_N \):

\[
P(x_N) := \prod_{i=0}^{n} P(x_i \mid x_{J(i)}). \tag{35}
\]

One should observe that from Equation (35) the \((X_G, \rightarrow)\) is a causal network compatible with the distribution \( P \) over \( X_G \); in other words, an organizational causal model over the GQM+ Strategies grid is constructed.

Using Equations (34), (32), and (33) will allow for the calculation of the missing PDT values, as explained in Section 8.3.4.
Causal Reasoning

The purpose of introducing the causality theory to the GQM* Strategies was to provide support for organizational causal reasoning, i.e. reasoning over the GQM* Strategies grids. Formal reasoning methods provide more sophisticated tools for analyzing grid structures. With the use of an organizational causal model, the same experts can update the probabilities (i.e. the probability distribution over a grid) when they become aware of the new data and new circumstances (context and assumptions) and obtain information about goal achievement.

Furthermore, the probability distribution function represents a quantification of the experts knowledge about the causal dependencies among different organizational goals. In other words, it is an alignment knowledge base. Having that alignment knowledge captured into the organizational causal model and using the formal causal reasoning mechanisms allow for a query of the knowledge base.

8.3.4 Obtaining Probabilities from Experts and Defining Joint Probability Distribution Function

In Section 8.2 the anatomy of the grid derivation process was analyzed. Defining a simple \( i \)-strategy is a unitary building block of the GQM* Strategies grid derivation process. In other words, the attention of the experts while they are developing grids is on simple \( i \)-strategies; therefore, the question regarding the effectiveness of those strategies is implicitly present because the experts are trying to design strategies that will deliver the desired effects, i.e. be effective.

Therefore, the questions for assessing the effectiveness of a simple \( i \)-strategy can be formulated as:

- \( Q \): What is the likelihood of having a successful simple \( i \)-strategy if all derived goals from that \( i \)-strategy are already achieved?

Sometimes it is more convenient to ask opposite question:

- \( Q_{neg} \): What is the likelihood of having an unsuccessful simple \( i \)-strategy if all derived goals from that \( i \)-strategy are already achieved?

Where the scale for both questions is given with nine-point discrete likelihood probability values. Due to the fact that there is no absolute certainty in this kind of
predictions, the probabilities of a certain event \((p = 1)\), and an impossible event \((p = 0)\) are excluded. Furthermore, the scale can be organized in a tabular form (Table 21) that has three major categories (LOW, MEDIUM, and HIGH) for assessing the likelihood and three self-assessing categories (optimistic, neutral, and pessimistic) that are used for the fine tuning of the expert’s assessment.

**Table 21. The scale for assessing probability distribution markers.**

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Optimistic</th>
<th>Neutral</th>
<th>Pessimistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOW</td>
<td>.3</td>
<td>.2</td>
<td>.1</td>
</tr>
<tr>
<td>MEDIUM</td>
<td>.6</td>
<td>.5</td>
<td>.4</td>
</tr>
<tr>
<td>HIGH</td>
<td>.9</td>
<td>.8</td>
<td>.7</td>
</tr>
</tbody>
</table>

The information provided by experts and values obtained from Table 21 characterize the probability distribution of a strategy, referred as the *probability distribution markers* and denoted as \(\mu(\sigma)\) for a simple strategy \(\sigma\).

Answer to the question \(Q\) represents what is defined as the measure of the strategy effectiveness in Section 8.3.3 or the answers formally provide information for Equations (31) as:

\[
P(\hat{\sigma}^j_i | x^*_i) = \mu(\hat{\sigma}^j_i) \quad \text{or} \quad P(\hat{\sigma}^j_i | x^*_i) = \mu(\hat{\sigma}^j_i) = 1 - \mu(\hat{\sigma}^j_i)
\]

where if the values from Table 21 are used for obtaining answer on question \(Q\), they represent probability distribution markers for having successful strategy \(\mu(\hat{\sigma}^j_i)\), while answers to opposite question \(Q_{neg}\) correspond to \(\mu(\hat{\sigma}^j_i)\).

In order to fully define a probability distribution function, all possible outcomes for a simple \(i\)-strategy have to be considered and a probability distribution table must be defined for it, where the outcomes of interest are "having successful strategy" and "strategy failed." Using the probability distribution marker \(\mu\) is sufficient to fully specify the probability distribution table (PDT) for a simple \(i\)-strategy as shown in Table 22.

**Table 22. A probability distribution table for a simple \(i\)-strategy**

<table>
<thead>
<tr>
<th>PDT for (\sigma)</th>
<th>Strategy (\sigma) successful</th>
<th>Strategy (\sigma) failed</th>
</tr>
</thead>
<tbody>
<tr>
<td>All derived goals achieved</td>
<td>(\mu)</td>
<td>(1 - \mu)</td>
</tr>
<tr>
<td>Some of derived goals not achieved</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

According to the simple \(i\)-strategy definition, the strategy can be effective, i.e. successful, only when all derived goals from that strategy are achieved; otherwise,
the strategy cannot deliver the expected effects. Therefore, two basic scenarios are possible: (1) when all derived goals are achieved and (2) when at least one of derived goals failed. The first scenario corresponds to the first row in Table 22, and those values are assessed by experts via probability distribution markers. In the second scenario, using the definition of the simple $i$-strategy, the values are predefined as $P(\sigma) = 0$ and $P(\sigma') = 1$ (Table 22).

In the addition to the empirical assessment of the effectiveness of simple $i$-strategies, it is necessary to assess the probabilities of bottom-level goals (i.e. leaf nodes in a tree structure). It is expected that bottom-level goals are specific enough, so an expert, after analyzing all the related context and assumption elements regarding the goal, can provide an answer to the following question:

- $Q_G$: What is the likelihood of having a bottom-level goal achieved?

The answer to the question is the probability distribution marker $\mu(G)$ with exactly the same scale as that of simple strategies (see Table 21). The outcomes of interest are "goal is achieved" and "goal is not achieved." Therefore, for each bottom-level goal, the probability distribution markers define PDT (Table 23).

Table 23. A probability distribution table for the bottom-level goals $G$.

<table>
<thead>
<tr>
<th>$G$ achieved</th>
<th>$G$ not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu(G)$</td>
<td>$1 - \mu(G)$</td>
</tr>
</tbody>
</table>

Having assessed the probabilities using probability distribution tables for the simple $i$-strategies and leaf goals, makes it possible to calculate the values for the distribution tables of the remaining goals in a grid. The entire process, i.e. steps, for defining probability distribution tables over the GQM+$^*$ Strategies grid can be summarized with following steps:

- **Step 1.** Assess probabilities of the bottom-level goals. By analyzing the context and assumptions of each bottom-level goal, assess how likely it is that the goal will be achieved. This will provide PDT values for the bottom-level goals (Table 23);
- **Step 2.** Assess the probabilities of all simple $i$-strategies in a grid. This will provide PDT values for the simple strategies in the form of Table 22;
- **Step 3.** Generate PDT for all goals in a grid from the experts’ assessments (Repeat for all goals in a grid):
  - (3.1) **IF a goal is a bottom-level goal THEN** this is already completed with Step 1.
– (3.2) **IF** a goal is addressed with a singleton simple strategy **THEN** PDT values for the strategy are also PDT values for the goal.

– (3.3) **IF** a goal is addressed with a composition of strategies (i.e. with a superstrategy) **THEN** depending on the decomposition types (AND or OR) using Equations (32) and (33), calculate PDT values for the goal.

Finally, the probability distribution tables and the structure of the causal model fully specify a joint probability distribution function over the grid in the form of Equation (35).

For example, Figure 35 shows defined values for probability distribution tables over a GQM+Strategies grid. Experts provided values for probability distribution markers (μ values for each simple i-strategy), and for bottom-level goals G3 and G5. Using the process explained above (Step 3.2), goals G0, G2, and G4 are addressed using singleton simple i-strategies; therefore, the PDT for those goals have the same values as PDT for corresponding strategies. In this micro example, the most illustrative is goal G1, because it is addressed with a compound strategy, i.e. two simple i-strategies σ1 and σ4. The probability distribution table for the goal G1 captures dependence relationships between G1 and a set of derived goals {G2, G4} at the next lower level. Depending on the logic of combining strategies σ1 and σ4 the resulting PDT for G1 can have different values, as shown in Figure 35.

**AND-type logic.** When both strategies are needed for the successful realization of goal G1, then the value for G1 achievement is calculated using Equations (34) and (32). In other words, the probability of having goal achieved is equal to the probability of having both strategies effective at the same time. Otherwise, if one strategy fails (i.e. one of the derived goals is not achieved) it will cause the failure of goal G1; therefore, the values in the PDT are set to zero and one.

**OR-type logic.** When alternative strategies for goal G1 are assessed, then the value for having G1 achieved is calculated using Equations (34) and (33). In other words, the probability of having not achieved the goal is equal to the probability of having both strategies fail at the same time. Only in the situation when both strategies fail, it is expected the goal G1 cannot be achieved; therefore, the values in the last row of the PDT are set to zero and one.
This example illustrates the effects of different strategy types in a grid. If the PDT values for OR-type and AND-type are compared, it is evident that OR-type probabilities for having a goal achieved are higher, which is one of the characteristics of the alternative strategies (Section 8.2.2); they tend to increase the probability of upper goals success, but this only occurs with the increased costs of redundant strategies. Another interesting observation is that AND-type strategies tend to experience reduced probabilities when they are combined due to the multiplication of probabilities. A practical explanation for this is that organizational units increase a set of favorable events that need to occur at the same time, which is accompanied by a larger set of undesired events that might occur. So, instead of having one unit that needs to achieve its goals, there are several units that need to achieve their goals at predefined timeframes, which also increases the probability of unexpected and unplanned events that might occur within these units.

### 8.4 Evaluation of the Solution

The solution is evaluated analytically and empirically. The analytical evaluation demonstrates the usefulness of the solution by showing how to define a query over the GQM+ Strategies grid that returns information about the size of the risk embedded
in a grid. As for the grid embedded risk, it is possible to define queries for testing relationships among different goals in a grid.

The goal of the empirical evaluation is to test the assumption that experts are capable of providing information regarding the effectiveness of the simple $i$-strategies. For that purpose, a small empirical study was designed and conducted using the GQM$^*$Strategies experts from industry and academia.

### 8.4.1 Analytical Evaluation: Analyzing GQM$^*$Strategies Grid Embedded Risk

Section 6.3 explains how to identify risky goals in a GQM$^*$Strategies grid. However, it does not discuss how to analyze the impact risky goals have on other goals in a grid structure. For example, this can be useful in a situation when several alternative grids are derived for the same set of top-level organizational goals. The information about the risk embedded in each grid can be used as one of the selection criteria, or it can be used as an overall risk measure for a grid when monitoring the grid execution. Risky goals spread a threat (i.e. a risk influence) onto other goals in a grid to some extent, assuming that the "size" of the risk influence can be quantified as the grid embedded risk. Such a risk measure can provide a better understanding of the influence of the risky goals on overall organizational success.

**Definition of the Grid Embedded Risk Measure** Let $\Xi_G = \langle \mathcal{X}_G, \sigma(G), \mathcal{M}_G \rangle$ be a GQM$^*$Strategies grid, and $G$ is a set of organizational goals partitioned into two sets: a set of risky goals $G_R$ for $R \subset N$, and a set of non-risky goals $G \backslash G_R$. Let $P$ be a joint probability function over a grid $\Xi_G$. The probability of having all the goals in a grid fail due to the failure of the risky goals is given by:

$$\mathbb{R}(\Xi_G) = P(\tilde{x}_{N \backslash R} | \tilde{x}_R) = \frac{P(\tilde{x}_N)}{P(\tilde{x}_R)}$$

and represents the grid embedded risk measure, where $P(\tilde{x}_R)$ is the probability that all risky goals fail, and that probability is larger than zero (otherwise these would not be considered risky), and $P(\tilde{x}_N)$ is a probability of having all goals in a grid fail.

The grid embedded risk measure, $\mathbb{R}$, queries the probability of a situation, i.e. the worst-case scenario of having the remaining goals fail if the risk threat became real and causes all risky goals to fail. The risk measure takes values from the $[0, 1]$ interval. With
the increase in the risk measure values, the impact of the risky goals increases. When the \( \mathcal{R} \) values are close to zero, the size of grid embedded risk is small. In the situation when \( \mathcal{R} = 1 \) the failure of risky goals will automatically cause the failure of all other goals in a grid. The logic of the grid embedded risk measure and apparatus introduced here will work on grids of any size.

Consider the following example scenarios. The example grid with the defined probability distribution function is given in Figure 35. Using techniques described in Chapter 6, the experts identified four different cases with different sets of risky goals, and using the joint probability function, the values for the grid embedded risk measure are calculated (Table 24). The cases can be seen as snapshots of the same grid in different time periods.

<table>
<thead>
<tr>
<th>Case</th>
<th>Set of risky goals ( G_k )</th>
<th>( \mathcal{R} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>( {G_5} )</td>
<td>0.3</td>
</tr>
<tr>
<td>2.</td>
<td>( {G_4} )</td>
<td>0.115</td>
</tr>
<tr>
<td>3.</td>
<td>( {G_4, G_5} )</td>
<td>0.3</td>
</tr>
<tr>
<td>4.</td>
<td>( {G_3, G_5} )</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Here, it is shown how to calculate the \( \mathcal{R} \) for case 2, where the goal \( G_4 \) is considered to be risky; the example grid is shown in Figure 35. According to the definition of the embedded risk measure (Equation 36) for case 2:

\[
\mathcal{R}_2 = \frac{P(\bar{x}_0, \bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4|\bar{x}_5)}{P(\bar{x}_5)}
\]

First, \( P(\bar{x}_5) \), the marginal probability of having goal \( G_5 \) fail, will be calculated. This can occur when \( G_4 \) fails given that \( G_5 \) is achieved, or when \( G_4 \) fails, given that \( G_5 \) is not achieved. Formally, this can be written as:

\[
P(\bar{x}_5) = P(\bar{x}_4|\bar{x}_5)^* P(\bar{x}_5) + P(\bar{x}_4|\bar{x}_5) P(\bar{x}_5) = 0.4 \times 0.8 + 1 \times 0.2 = 0.52
\]

where the values for the probabilities are taken from the PDT example given in Figure 35.

Now, remains to calculate \( P(\bar{x}_0, \bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5) \), the probability of having all goals in a grid fail. Knowing the grid structure, i.e. the dependencies among random variables, the factorization rule (Equation 28) can be used and written as follows:

\[
P(\bar{x}_0, \bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{x}_5) = P(\bar{x}_0|\bar{x}_1) P(\bar{x}_1|\bar{x}_2) P(\bar{x}_2|\bar{x}_3) P(\bar{x}_3|\bar{x}_4) P(\bar{x}_4|\bar{x}_5) P(\bar{x}_5) = 0.06
\]
For case 2 in Table 24, the grid embedded risk measure is stated as:

$$R_2 = \frac{0.06}{0.52} = 0.115.$$  

With respect to the contained risk in the grid, case 2 is the best case of all cases because the $R_2$ has the smallest value. In other words, the threat introduced by risky goals is the smallest. According to Table 24, case 4 is the worst case because the risk measure has the highest possible value $R_4 = 1$. In this case, all bottom-level goals are also risky goals (Figure 35); therefore, a failure of these goals will consequentially mean a failure of the upper-level goals and strategies.

Cases 1 and 3 are the same with respect to the grid embedded risk ($R_1 = R_3 = 0.3$). Interestingly, the set of risky goals in case 1 represents a subset of the risky goals in case 3, but they still introduce the same size of the grid embedded risk. For this particular situation, the explanation lies in the fact that goals $G_5$ and $G_4$ form an isolated sub-grid (Figure 35), where $G_5$ is a bottom goal, and it, alone, impacts only the goal $G_4$.

### 8.4.2 Empirical Evaluation

In Section 8.2.3, the question regarding the effectiveness of a simple $i$-strategy was identified as the main question that concerns experts while they are developing the GQM$^+$Strategies grid. Because of this, it was assumed that experts are capable of providing feedback (expectation or prediction) regarding the effectiveness of the simple $i$-strategies that they defined. Therefore, the main objective of this empirical evaluation is to test the assumption used to develop/design the solution:

**Experts are capable of providing information regarding the effectiveness of simple $i$-strategies.**

In other words, this empirical study focused on testing the capability of experts to provide certain information; therefore, the experts were subjects. Experts were practitioners who already participated in the GQM$^+$Strategies grid derivation; either it was a pilot for their organization, or it was an exercise. For this study, it is important that experts had grids that they developed or participated in developing.

The study approach was to design small exercises and data collection protocol in the form of surveys and interviews. The exercises were performed on the selected pieces of the GQM$^+$Strategies grids. The exercise was about assessing the effectiveness of strategies via strategy assessment cards. The design of the study was simple enough that
it can be conducted remotely (e.g. via Skype and email) and in person. These were the three main activities of the study:

- **Study preparation** involves selecting experts and grids, sampling simple i-strategies from selected grids, and preparing strategy assessment cards;
- **Conduct exercise and collect data** from experts who are involved with strategy assessment via a survey about the quality of assessment and an interview discussion about the assessment process;
- **Data analysis** is mainly focused on findings from the interviews and assessment follow-up survey.

Each of these activities is explained in the following sections.

**Study preparation**

The first step was to identify and select experts. Because the number of the GQM* Strategies applications is relatively small (less than 10), there was no elaborate strategy for selecting experts. Experts were also selected from the academia in addition to the industry. The experts from academia were either involved with developing the GQM* Strategies approach or they had experiences with industrial applications of the approach. In total 8 experts from academia were approached, and of those approached 3 volunteered to participate. The primary purpose of having study with experts from the academia was to test run the study before proceeding with industry experts. Industrial experts were selected among those participants who had the best possible understanding about the approach, i.e. they received the GQM* Strategies training and were involved in the grid derivation process from the beginning (they are also referred as the method champions). Although during the years of dissertation work some 60 people were involved with the GQM* Strategies grid development, for this particular study the available access was to only two companies that were involved with the ongoing research projects. And, in those two companies 3 experts were identified as method champions, they were approached and all of them accepted to participate in the study. After obtaining the experts’ commitment, the potential grids were known. If an expert participated in developing more than one grid, the most current grid was selected.

After selecting experts and grids, the next step was to selected simple i-strategies. Table 25 shows a potential sampling strategy for selecting simple strategies from a grid. The experts did not conduct the exercises for the entire grid. For an exercise of this
Table 25. Sampling strategies.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Convenience sampling</td>
<td>Pick any number of simple strategies.</td>
</tr>
<tr>
<td>B. Criteria-based sampling</td>
<td>Select strategies that satisfy the following criteria:</td>
</tr>
<tr>
<td></td>
<td>1. Select equally simple strategies from different grid levels (e.g. top, mid, and bottom levels)</td>
</tr>
<tr>
<td></td>
<td>2. If possible, include a number of non-singleton simple strategies</td>
</tr>
<tr>
<td></td>
<td>3. Include simple strategies that experts consider the most critical</td>
</tr>
<tr>
<td>C. Random sampling</td>
<td>Enumerate all simple strategies in a grid, and then use a random number generator to pick a subset of strategies.</td>
</tr>
</tbody>
</table>

scope, a significant time and effort is required from an expert. Therefore, pieces of a grid were sampled and used for the assessment exercise, which helped to reduce the costs of study and ensured a high response rate from the industry experts.

For each selected simple strategy, an assessment card was prepared (Figure 36 shows an example of the assessment card). The assessment card had two sides. On the left side, a single i-strategy was presented with the related goal and all refined goals with the strategy. In addition, all context and assumption elements that were relevant for presented strategy and goals were given.

Fig 36. Strategy assessment card.

The right side of the card (Figure 36), was used to collect the feedback from an expert regarding the strategy effectiveness. The first question was about the likelihood of
the effectiveness of the given strategy if all derived goals were already achieved. The second question was used to fine-tune the answer from the first question. The scales for both questions were designed to conveniently obtain probability distribution markers from Table 21.

**Conduct Exercise and Collect Data**

At the beginning of the exercise, the experts were familiarized with the objectives of the exercise, what was expected from them, and how the data would be collected. After that, the assessment cards for selected strategies were presented to the experts, and they were asked to assess the effectiveness of those strategies, followed by a short survey for evaluating the quality of the assessment. The survey questions are given in Appendix 3.

Finally, the exercise was concluded with an interview discussion about the assessment process. The semi-structured interview protocol allowed an exploration of more detailed issues during the assessment. The interview protocol is given in Appendix 4.

The data were collected using different means. First, the assessment cards were collected; however, due to confidentiality reasons these raw data in connection with specific organizational strategies and goals are not disclosed. Therefore, the use of surveys and interviews allowed data about the assessment process to be obtained without disclosing confidential company data.

**Analyze Data**

Data analysis had two parts. First, the assessment cards were analyzed to identify any negative answers, i.e. *I am not able to provide that kind of information*. The second part of analysis focused on survey and interview data to understand why experts were (or were not) capable of providing an assessment about the strategy effectiveness, which was the objective of this study.

**8.4.3 The Study Setting and Findings**

The empirical evaluation involved six experts (subjects) in total. Of those six, two were from academia (33%), and the rest 67% were industry representatives. The total number of assessment cards was 17 (2.8 cards per expert), and all the strategies were picked using a convenient sampling strategy from the chosen grids. Only two exercises were
conducted over Skype and email, and others were conducted in person. The purpose of involving experts from academia was to conduct a test-run of the study design before proceeding with industry. They were assessing strategies form a grid that was used for developing the GQM+Strategies approach.

The interviews were not recorded; only notes were taken during the interviews, and summaries were written after the interviews. A summary of interview notes is given in Table 31 in Appendix 5.

The results of follow-up survey are given in Figure 37. Each bar chart in the figure corresponds to one question from the survey. The questions concern (1) the difficulty to provide a strategy assessment, (2) interviewees familiarity with context, and (3) the assessment approach, i.e. what was the mental process during the assessment. The results of the survey and interviews are discussed in the following sections.

**Feasibility of the Strategy Assessment**

One of the main objectives of the empirical study was to investigate if experts can provide an assessment or prediction about the effectiveness of a simple i-strategy. The overall feedback is that this kind of assessment is feasible (5 of 6 experts). However, one expert responded that it is extremely challenging to provide such feedback (Figure 37). Later, during the interview discussion, the reasons for this response were discovered. The expert in question was in academia, and he felt that the hypothetical example he had developed was not realistic enough to provide such assessments (i.e. the assessments will be hypothetical as well). Therefore, it was interesting to compare his statements...
with the second expert who was also from academia. The second expert was confident about providing assessments for the same grid example, but he noted that when the environment has been abstracted with context and assumption elements, it was easy to make the assessment, and he commented:

“ If I would need to think about specific individuals without abstracting it might turn out to be more challenging.”

The experts from industry were in consent that providing strategy assessment is certainly achievable. The only variation was with respect to providing an assessment for strategies that they are directly involved with or for some strategies coming from other organizational units. In the first case, the quality of assessment was the highest. Obviously, the familiarity with environment is of the highest importance for providing reliable assessments.

It was pointed out that time is also an important factor for the assessment. For example, one assessment exercise took place six months after the grid development, and the strategies and goals were in the process of being implemented, so the expert was more confident with his assessment because that time period was sufficient to see if original predictions and assumptions were correct.

**Familiarity with the Environment (Context)**

In order to provide good assessments an expert have to be familiar with goals and strategies that he is assessing, but he must also be familiar with the environment. Many factors that can impact strategy effectiveness are residing in the business environment. Therefore, the second question the experts were asked was about their familiarity with context. Of the subjects, 84% said that they had sufficient knowledge about the environment to make valid strategy assessment (Figure 37).

During the interview, a question how familiar should someone be with the context in order to provide good assessment was discussed. That discussion can be best summarized as follows:

Familiarity with context and environment helps to make better assessments.

In an ideal situation, the information captured by the grid (i.e. presented in an assessment card) should be sufficient for an outsider [someone who was not involved with the grid development] to assess the strategies. However, if someone is familiar with the context, he or she will be able
to provide an equally good assessment with less information presented in the assessment card. It is realistic to expect that GQM + Strategies grids capture just enough information that someone who is familiar with the business domain and market can provide good strategy assessment.

Familiarity with the context is a critical factor for the assessment. People who are familiar with context are able to see environment in the abstract and focus their attention on the most important issues.

**Expert's Assessment Approach**

Experts were also asked to explain their approach, i.e. mental process, for strategy assessment. Understanding the approach can help to improve the assessment process itself. Of the participants, 84% used a straightforward approach to analyzing strategy and related goals for the given context and assumption elements, and then they deduced the assessment.

However, one expert claimed to have a different approach, which he explained during the interview. His approach is summarized below.

First, I analyzed the information I was given in the assessment cards. I checked if it was appropriate information (e.g. if it was too high-level, or was too abstract), and it was all OK, except one strategy that looked to be at a level that was a bit too high. After that, I compared the presented model [goal, strategies and context/assumptions] with the model in my head, i.e. my understanding and knowledge about that particular goal/strategy. The presented model fit well with the model in my head; otherwise, I would have thought about why the presented model did not fit the model in my head.

It seems that the most common approach (used by all experts) is a comparison of implicit and explicit models, where the explicit model of goals and strategies is documented using the GQM + Strategies grid while implicit models represent experts’ knowledge, experience, and beliefs about certain strategies. The only variation is with respect to the comparison criteria that they themselves determined.

All interviews were concluded through a short discussion of the usefulness of the strategy assessment, identifying if the assessment should be conducted for all strategies in the GQM + Strategies grid. The points raised include (see Appendix 5):
– **Strategy assessment is important and necessary while you are chaining the grid.** The assessment can help identify the strategies need to be changed and changes that need to be evaluated, i.e. if they are increasing the strategy effectiveness or not.

– **A tool for checking the quality of strategies and goals.** Here, the quality defines an intrinsic value of the strategy. Good strategies should have higher chances of achieving goals than strategies of lower quality.

– **Can show missing or insufficient information.** The assessment process is a form of reflective thinking about previously defined strategies, and it can identify missing information, e.g. a strategy assessed as ineffective can be an indication of missing key assumptions about the strategy and related goals.

– **Can generate discussion points for strategy review meetings.** Assessment results can be used to start and focus discussion on the most critical elements in the grid.

– **It can be used to indicate the level of agreement toward specific strategies.** When the assessment is conducted by a group of people, it can provide a quick indication of the group’s opinion toward the joint strategies.

– **Checking the belief of the wider group of people in an organization.** This is similar to the previous point, but with one difference. The purpose would be to ask wider group of people (the people who are supposed to work on the implementation of strategies) and identify their beliefs. If people do not believe, then it will be difficult to enforce strategies in the implementation phase.

With respect to the validity of the empirical study, the sample was representative because a large number of organizations that are using the GQM+ Strategies approach for different purposes (i.e. population) participated. The objectives of the study were clear, and direct questions were asked. The participants were able to complete the assessment exercise, which is a proof of feasibility. Furthermore, a short survey and interviews were used as instruments to collect data about the assessment process. As an overall conclusion, the study findings support the initial assumption about experts’ abilities to provide predictions about simple i-strategy effectiveness.

The practitioners found several potential applications of the strategy assessment process. Therefore, the assessment is not just feasible, it was also perceived to be useful. Furthermore, the analytical evaluation of the solution shows how the information obtained through assessment can be further used to analyze the relationships among different sets of goals, particular, the impact of risky goals onto the successfulness of the remaining goals in a grid via measuring the grid embedded risk.
IV Conclusions
9 Results and Conclusions

The main goal of this dissertation was to evolve the GQM+Strategies approach by providing an organization with the capability to:

1. Apply the work of value-based software engineering to directly address the return-on-investment (ROI) of their goals and strategies via an evaluation of the costs and benefits of the goals and strategies that were chosen;
2. Calculate a set of earned value metrics that will allow organizations to effectively monitor the implementation of the organizational goals and strategies with respect to costs, schedules, and benefits realization;
3. Identify the risks associated with not achieving various sub-goals in a grid by analyzing goal risk exposures and acceptable risk levels for the estimated goals’ cost-benefit ratio; and
4. Assess the threats of risky goals on other goals in the grid via an evaluation of goals dependencies using the formal goal-strategies-goals models.

These organizational capabilities can be accomplished by extending the existing GQM+Strategies grid with additional elements for monitoring the organizational value creation process and constructing organizational causal models. The additional elements that address the value creation process are compliant with value-based software engineering (VBSE); therefore, organizations can benefit from the VBSE practices and tools as well. Organizations that are using or planning to adopt the VBSE practices can use the value-based version of the GQM+Strategies as a notation for their VBSE activities. A major benefit is that organizations will be able to analyze dependencies between organizational goals and strategies. Such analysis brings additional understanding and insights to how different organizational units relate, i.e. align. This is achieved by introducing causality theory to the GQM+Strategies approach. This provides an organizational causal model that aligns different organizational units and provides support for implementing the value-based software engineering activities, i.e. managing the organizational value creation process. These and other contributions are elaborated in the Section 9.1.

Dealing with an organizational approach poses challenges with respect to the empirical evaluation of the entire solution. The validity issues and threats are discussed in Section 9.2. Finally, future work is discussed in Section 9.3.
9.1 Contributions

The work presented in this dissertation complements and extends the current state of practice and theory as discussed in the following.

The current GQM+ Strategies approach (Section 2.3) views the definition of strategies as a selection decision, i.e. select the most promising strategies for a goal [13]. This process of strategies selection is now further elaborated and clarified by the introduction of causality theory, as discussed in Chapter 8. Making the part of the GQM+ Strategies grid, i.e. that the goal hierarchy, explicitly represent an organizational causal model, leads to a new cognition about the grid derivation process. As the result, experts that are developing grids can define strategies as causal relations and quantify their beliefs about the causal effects in terms of probabilities. Furthermore, as shown in Section 8.4.3, using the strategy effectiveness assessment acts as a beneficial tool for deriving the GQM+ Strategies grid. Because, the assessment questions can be used as tools for building a consensus among goal owners, especially with large groups, discussing the assessment findings can help identify alternative or better strategies.

The solution presented here is fully aligned with the VBSE concepts (Section 2.4). The benefits realization analysis is carried out while analyzing the context of the goals that are documented with the new goal template with the added value dimensions (Section 6.3). The process of defining goals represents stakeholder value proposition elicitation and reconciliation. Refining organizational goals with strategies and documenting the relevant context/assumption elements is a way of conducting business case analysis. Identification of the critical GQM+ Strategies sub-grid and regular context and assumption updates leads to continuous risk and opportunity management. Value-based monitoring and control is supported by organizational earned-value metrics. The GQM+ Strategies grid structure enables users to act on changes by selecting the best possible opportunity, i.e., viewing change as an opportunity.

Boehm [37] pointed out that it is important to understand the links among technical decisions, context, and value creation in different situations in order to improve decision making. Furthermore, dynamic monitoring and control mechanisms take into account that linkages and different sources of value are necessary to guide decision makers [37]. The GQM+ Strategies grid structure helps users understand the relationship between context and the value creation process more clearly. Documenting goal’s strategies elements capture relevant information about a particular situation and offer an opportunity
to study value-based decisions and actions for that situation. These studies could be a part of the organizational learning process.

The utilization of GQM+ Strategies for business value analysis addresses management blind spots (linkage, reach, people, and time) discussed in Section 2.4 in the following ways: Linkage—the GQM+ Strategies grid explicitly exposes all relationships among grid elements within and among different organizational levels; Reach—the grid derivation process permeates the entire organizational structure, from the top-level to the operational levels, and makes the breadth of change visible; People—the grid derivation process is transparent and requires the involvement of people from all organizational levels, as those very same people are defining their own goals and strategies to address upper-level goals and strategies, which is a preparation and motivation for change; And Time—one of the business goal’s dimensions is time (timeframe), which is explicitly part of the process of defining and analyzing business context dynamics and value while defining goals and strategies.

Mapping of the GQM+ Strategies approach onto the VBSE framework identifies theories that are also relevant to the approach. The theoretical framework used to develop VBSE key activities (referred as “4+1” theories, Section 2.4.1) can be applied to the GQM+ Strategies approach. If the GQM+ Strategies is positioned into the center of the “4+1” theoretical framework, all theories and connections make sense from the perspective of the GQM+ Strategies approach. First, developing grids is about gaining a consensus and agreement about organizational goals and strategies, which includes defining goals that everybody can contribute to with the clear understanding of benefits that those goals bring to them, which is a way of creating win-win situations (i.e. Theory W and Utility Theory). The GQM+ Strategies grid captures links and relationships among different goals and strategies. Those relationships model dependencies (Chapter 8) among different organizational units (or groups within units), which are issues that deal with Dependency Theory. Documenting value propositions with the GQM+ Strategies goal hierarchy, analyzing costs and benefits, and identifying associated risks falls under Utility Theory. Goals and strategies should guide operational activities by providing a multi-level organizational decision-making framework, i.e. Decision Theory. Finally, the whole idea of aligning measurement schemes with organizational goals is about having better control structures and mechanisms when implementing organizational strategies (Control Theory). This fitness of the “4+1” theoretical framework for the GQM+ Strategies approach is another argument that these two approaches support each other well.
The solution presented here is the first approach that defines a unique structure (i.e., grid) to support all VBSE key activities. The VBSE contributions are contained within different subareas that correspond to the VBSE activities [113]. However, contributions that sought to provide support for all VBSE activities were not found in literature (Section 2.4).

The contributions to the areas of the software measurement alignment (Section 2.2) and value-based software metrics (Section 2.4.3) follow. The software measurement alignment area benefits from the formal organizational causal model (Chapter 8). In addition to integrating and aligning the GQM measurement schemes, it provides an interpretation framework that can interpret lower-level metrics (that are measuring the achievement of the lower-level goals) in the context of the top-level goals, by showing how much progress that was made at lower levels affects opportunities of achieving the top-level goals.

Value-based software metrics are not adequately covered by the existing literature [113]. The main reason is that value-based metrics are context dependent, and in some cases value is not measurable as a simple scalar quantity [47]. Therefore, Boehm & Sullivan suggested the development of a value-based version of GQM in conjunction with the Experience Factory [47, p.7] in order to manage value creation at the organization level. The GQM*Strategies approach, with the value-based extensions developed in this dissertation, achieve exactly that, by defining organizational earned-value metrics (Chapter 7) over the grid structure that enunciates the organizational value proposition, as discussed in Chapter 6. Metrics defined in this way are not just capable of monitoring the cost side, as traditional earned value metrics do, they monitor benefits realization as the organizational goals and strategies are implemented.

Monitoring and controlling the risks involved with strategy and goal implementation is important for achieving a success. The solution developed here adds new capabilities to the GQM*Strategies approach with respect to risk handling. First, the solution developed in Section 6.4 defines a means of identifying risky goals, and based on that, selected a sub-grid that needs further attention, resulting in the GQM*Strategies critical sub-grid (Section 6.4.2). The work in Chapter 8 introduced more formal ways of representing and analyzing grid structures, which enabled an assessment of the overall risk embedded in the grid (Section 8.4.1). This gives an organization the ability to make better decisions when considering alternative sub-grids or when checking the progress of a grid execution, i.e. monitoring how the overall risk in a grid is changing over time.
All the contributions are the result of developing a solution to answer the research questions (Section 4.6). In what follows, the research questions are revisited and the dissertation work is summarized. After that, the validity issues are discussed.

**RQ1: How can the GQM*Strategies approach be used to document value propositions across an organization?** The main challenges behind the research question (Section 4.6.1) were the existence of the multiple value propositions across an organization, coming from various groups of people and units and a need to consolidate and synchronize (i.e. align) all those value propositions into one organizational value proposition. Chapter 6 argues that organizational goals are defined according to the stakeholders’ perception of value and explains how GQM* Strategies is used to enunciate an organizational value proposition while developing an organizational goal hierarchy. The main features of the proposed solution are:

- Formalizing individual value proposition statements with the GQM* Strategies goal template (Section 6.2); the template is extended to incorporate the cost and benefits estimates of goals (Section 6.3). Quantifying costs and benefits of organizational goals helps to make the contributions of individual value propositions more explicit to an organization;
- Goals from various organizational levels and units are part of an organizational goal hierarchy. The hierarchy aligns value propositions represented by organizational goals and integrates them into one organizational value proposition (Section 6.2.1).
- The grid derivation process involves people (i.e. goal owners or stakeholders) from various organizational levels, and communicates what is valuable for an organization at those levels;
- Using the GQM measurement approach, a set of metrics is defined for evaluating the value propositions (i.e. organizational goals). In other words, the value propositions are not only formalized by the goal template, they are measurable, which is important for managing the value creation process.

Considering all previous points, the GQM* Strategies grid is a structure that captures organizational value by providing **bounded flexibility** (Section 6.2.1) to goal owners, so they can define goals according to their best knowledge of what would be valuable for their organization or unit. The goal owners’ perception of value is based on their knowledge and experience, therefore the entire GQM* Strategies grid derivation process
can be seen as a formalization and dissemination of that specific knowledge. That knowledge helps to *align* different parts of an organization.

Also, integration of the costs and benefits estimates with the grid structure provides a basis for defining ROI models for sets of organizational goals and strategies.

**RQ2: How can the GQM* Strategies be instrumented to monitor and control a value creation process?** The value materialization process in software-intensive organizations is complex, and to a great deal, intangible. The main concerns behind the research question are how to monitor the dynamics of the value creation process, including the issues of tracking actual costs and benefits and monitoring the value realization risks (Section 4.6.2). The GQM* Strategies based solution presented in this dissertation helps manage the value creation process by:

- Connecting an organizational value proposition with the GQM* Strategies grid structure and providing an initial set of metrics, as discussed for RQ1, which are used as baselines for monitoring the progress of implementing organizational goals and strategies;
- The actual costs and benefits are collected through the cost–benefit GQM graph attached to the organizational goal hierarchy (Section 7.2).
- Value creation occurs during the implementation (i.e. execution) of the GQM* Strategies grid. The set of basic and derived organizational earned value metrics helps monitor the dynamics of the grid execution (Section 7.3). The organizational earned value metrics control both the costs and benefits realization of goals and strategies, which helps goal owners reach better conclusions about the value realization;
- Risk handling starts by analyzing individual organizational goals based on how much *risk* they carry (Section 6.4.1), and if that risk exceeds the *acceptable risk* level for the goal’s investment size and expected benefits (Section 6.4.1), then the goal is classified as a *risky goal* (Section 6.4.2)
- A set of risky goals in a grid defines the GQM* Strategies *critical sub-grid* (Section 6.4.2). The critical sub-grid helps direct the attention of managers during grid execution to the activities whose outcomes are important for the success of the organizational goals in the grid.

Obviously, the RQ2 is all about the dynamics of the GQM* Strategies grid execution. The grid itself represents an organizational value proposition that aligns the entire organization. During the execution of a grid, different parts of an organization can
perform differently, i.e. some will be more successful than others in achieving goals, and even some units can become misaligned. Such situations are identified with risky goals. For example, failing to achieve risky organizational goals can have serious consequences with respect to successfully implementing the entire grid. The grid embedded risk measure (Section 8.4.1) helps in the analysis and understanding of the seriousness of these situations.

The organizational earned value metrics defined over the GQM*Strategies grid provide a sufficient system for monitoring the implementation of goals and strategies with respect to budget, schedule, and benefits. Furthermore, after detecting the plan changes or deviations, the new circumstances can be effectively analyzed with respect to the context, assumptions, and risks involved. This leads to decisions that impact the value creation process in the best possible way for the given set of constraints.

**RQ3: How can the GQM*Strategies be used to analyze dependences between different value propositions across an organization?**

In other words, the issue addressed by this research question was how to quantify the goals-strategies dependencies in the grids (Section 4.6.3). Chapter 8 is dedicated to solving this problem by constructing the goals-strategies dependencies models. The construction of such models includes:

- Analyzing different patterns of goal refinement and identifying different strategy types (Section 8.2.2)
- Collecting feedback from practitioners (experts) about certain type of strategies, i.e. simple strategies, and goals, such as bottom-level goals (Section 8.3.4)
- Using the feedback provided by experts to calculate the contributions of lower-level goals to upper-level goals (Section 8.3.3)
- The formal reasoning structures to query the dependencies among different value propositions (Section 8.3.3)

The goals-strategies dependency model is defined over the GQM*Strategies grid; it quantifies the influences that different organizational units have on each other and on their value propositions. The most important benefit of the dependency models is the extraction and packaging the expert knowledge and predictions about certain relationships. In fact, it is these very same relationships that determine the alignment of an organization. Having that knowledge immersed into a model establishes possibilities for using the knowledge in different forms of queries. For example, Section 8.4.1
illustrates how to query the probability of failing to achieve non-risky goals in a grid if the risky goals in the grid are not achieved.

9.2 Validity Issues and Solution Limitations

Before discussing validity issues and the limitations of the solution, it is important to make the constraints that define and bound the research context explicit. The research context is the interface between researchers and the real world. The major constraints of the research context were:

– **Partial access to company data.** For these kinds of studies it is not realistic to expect to gain full access to company-specific data, despite having the confidentiality procedures in place; especially information about a company’s strategies and goals that are related to new technologies or products. This kind of information is very sensitive in the software and ICT industry. Therefore, the access to data is limited and partial, in other words only some goals and strategies are disclosed to researchers;

– **Newness of the related research.** The GQM* Strategies approach itself is a new methodology, consequently, there is not an established pool of companies that use the approach routinely on their own for a long time period. So, it is not possible to conduct a series of empirical studies, even at the level of quasi-experiments. Empirical research options are best when one can work with practitioners, gather feedback, and assessments of the work.

– **Research timeframe constraints.** Let assume that a company is using the GQM* Strategies approach on its own and is willing to use, i.e. test, the proposed solutions. The grid contains goals at different levels, and usually, the timeframes for the top-level goals are measured in years. Conducting a longitudinal study to evaluate an application of the solution exceeds the timeframe of a PhD dissertation.

If a solution is not further evolved to satisfy the needs of practitioners, it will never be adopted and used routinely. Therefore, it is important to maximize gains from each opportunity to work with practitioners and conduct different, small empirical studies. The purpose of those empirical studies is to collect feedback and supportive arguments, whether reasonable or not. These studies also identify where the research needs to be improved or changed. This can be used in conjunction with analytical paradigms to further evolve the solution.

In the following sections, the solution validity issues are discussed.
9.2.1 Internal Validity

The solution validity, i.e. internal validity, is concerned with the solution goodness or how well the solution is constructed and its usefulness, i.e. does it address relevant practical problems.

The dissertation claims that the GQM+Strategies’ extensions presented here address relevant problems, and consequently, the solution satisfies the usefulness criterion. The GQM+Strategies approach was piloted with four different organizations, meaning that partial grids were developed for those organizations, and additional tutorial and presentations of the approach were conducted on two other occasions (Section 4.2). Together, the research involved more than 60 participants. Qualitative methods were used, mainly interviews and participant observation techniques to provide the richness of data. Furthermore, a comparative analysis of two cases was conducted (Section 4.4) to support claims from informal feedback and to increase the validity of the findings. The two pilot cases selected had an equally high commitment from practitioners, meaning that the GQM+Strategies grids were developed to the same level of details and size. The pilot applications of GQM+Strategies were overseen by two different groups of researchers, which is important during the analysis phase to distinguish issues that are caused by the technology transfer process or by the approach itself. The analysis of feedback showed that there were no contradicting opinions about the most common issues or suggestions for further improvement of the approach. The feedback was used to formulate research question (Section 4.6), which represent a high-level requirements for the new solution.

The solution development is a process, and the best way to discuss the appropriate-ness of the solution is to make that process transparent and criticizable. For that purpose, the design research framework (Section 3) was adopted and used. The rationales for each activity or phase in developing the solution are summarized using the design hypotheses (Section 5.2), where outputs of each activity were evaluated analytically or empirically. The solution development activities represent Chapters 6, 7, and 8; where each chapter has a solution evaluation section.

The first activity (Chapter 6) involved including the value-based components into the GQM+Strategies meta-model in order to support monitoring and control of the value creation process. The extensions developed for that purpose were analytically evaluated (Section 6.5) through a conceptual analysis of the solution and the value-based software engineering framework. The VBSE framework predefines criteria and activities
that need to be performed in order to include the value perspective into the software development process. The results of the analysis showed that the solution is compliant with the exception of one key VBSE activity—value-based monitoring and control (Section 6.5.2). Therefore, the next activity (Chapter 7) focused on developing adequate metrics support for value-based monitoring and control, which resulted in organizational earned-value metrics defined over the GQM+ Strategies grid. The results were evaluated analytically, showing the compliance of the organizational earned-values system based on GQM+ Strategies with requirements of value-based monitoring and control (Section 7.4). Furthermore, the organizational earned-value system is not just complying with the value-based monitoring and control requirements, it improves the value realization feedback process itself through the introduction of the benefits-related earned-value metrics.

The last activity (Chapter 8) incorporated a causality framework to analyze the dependencies between goals and strategies in the GQM+ Strategies grid. The analytical paradigm for this activity involved formal mathematical modeling for constructing organizational causal models. The formalism that was used, did not just help solve the specific problem (i.e. quantifying dependencies among goals in a grid), it built a new theory that can be used to analyze a much wider spectrum of situations that might occur during a grid development. The solution was evaluated analytically and empirically (Section 8.4). The purpose of the analytical evaluation was to demonstrate the usefulness of the mathematical model, by explaining how to measure the grid-embedded risk (Section 8.4.1). The solution is based on a formal mathematical model, but it requires some interaction with people to provide inputs. The ability of people to provide those inputs in a feasible way is an important condition for the successful adoption of the solution. Therefore, the empirical evaluation tested that aspect of the solution (Section 8.4.2). The evaluation involved six experts, from which two of them were from academia. Considering the fact that there are six organizations that have developed GQM+ Strategies grids (Figure 21), this is about the maximum number of experts that could be obtained. The majority of experts (86%) were able to perform strategy assessments and found strategy assessment feasible (Section 8.4.3). Only one expert from academia felt uncomfortable about providing an assessment for his example grid. The experts from industry provided assessment for their grids, and they characterized the difficulty level of the assessment as average, compared to their daily work. Furthermore, the practitioners found several potential applications of the strategy assessment process, such as a tool for checking quality of strategies and goals or a tool for indicating the
level of agreement toward specific strategies. Therefore, the assessment is not just feasible, it was perceived as useful.

### 9.2.2 External Validity

The *external validity* is concerned with the generalizability of the solution, i.e. how well can it be applied and used in contexts other than research context. The major obstacles to solution generalizability are solution limitations. For the solution and given constraints, as discussed above, it is hard to give exact answers about the limitations, but the discussion itself is important because it shares concerns that can be the subject of future studies.

There may be limitations to the method itself and to the environment in which it was applied.

#### Limitations to the method

The major concern regarding solution generalizability is the *scalability* of the solution. Scaling up the solution, i.e. involving a larger number of organizational units, result in larger GQM+Strategies grids. The *size* and *complexity* of the grid are both problems with respect to scale-up. Complexity means the number of interrelated elements and their relationships (e.g. complex strategies for multiple goals). Maintaining those relationships requires more effort for larger grids than for small, simple ones. In fact, for small, simple ones it might be possible for experts to keep all those relationships in their heads, and to analyze them when needed. With increasing complexity, it is not possible to keep all those relationships in one’s head. Which leads to the questions: *what is the manageable size of the grid for an average expert, and how can grids larger than that size be managed in a feasible way?* Software tool support can help manage a complex grid by providing aids for defining, updating, and archiving grids for later retrieval (reuse), but, even with a software tool, there will be some limitations. The formal theory developed in Chapter 8 helps address the scalability issue to some extent. First, it allows experts to work with sub-grids (the grids that are of smaller and manageable size) and to quantify their expectations and beliefs for those grids. Secondly, using the causality framework those sub-grids can be integrated into larger grids, and using causal reasoning allows the definition of various queries to extract relevant information from the whole, integrated grid.
The second concern regarding method applicability is the cost of developing and maintaining the grid. It is evident that developing grids requires significant human resources, i.e. effort provided by experts, and this is probably the largest single cost factor for an organization. The question is: how do the grid costs behave when the grid complexity is increased? The encouraging finding is that for smaller grids developed during industry pilots, the practitioners recognized the benefits of developing grids, and they felt that the cost-benefit factor was satisfactory. However, if an organization in the process of adopting the approach does not recognize the benefits, the adoption will be terminated. Therefore, it is important to develop the cost-benefit models for the adoption of the GQM* Strategies approach. Such models will also help plan the use of adequate resources for the adoption process, and having more realistic planning increases the probability of successful adoption of the GQM* Strategies.

Limitations of the environment

Regarding the limitations of the application environment, the major obstacles are in connection with an organizational way of working. Organizations have established ways of doing business, and they have different practices, organizational policies, and even, rituals that constitute their way of doing business. Depending on their current way of working, the adoption of the GQM* Strategies based solutions can be more or less successful, or even a failure. Especially, when the GQM* Strategies is scaled up to the corporate level, and at that point organizational issues, like culture and resistance to change become important factors. For example, organizations that have established goal-driven, collaborative culture will be able to adopt the GQM* Strategies approach much easier than autocratic organizations that are nurturing military style chain-of-command leadership. The interesting question for further investigation is what organizational cultural models have positive effects on the GQM* Strategies adoption process? A particularly interesting aspect is the socio-organizational factors. In the software industry, the major assets are knowledgeable and creative people, and therefore, it is necessary to explore socio-organizational factors.

Another issue that is related to an organization’s way of working, but that is much easier to deal with than organizational culture, is the organizational set of skills. The GQM* Strategies approach, with all its extensions, adds new skills and capabilities to the organizational skill set. However, organizations need to have some skills in place before they can adopt and fully benefits from the GQM* Strategies and solutions presented in
this dissertation. For example, organizations should have basic skills and knowledge about GQM measurements before advancing to the GQM+ Strategies, and after they have capabilities to develop and maintain the grid, they would be able to use organizational earned-value metrics to manage their value-creation process. In other words, in order for such a solution to be successfully applied, it requires certain organizational capabilities, and if those capabilities are lacking, then the adoption in that environment will be less successful. This leads to a question of developing a capability model that will include a repository of relevant skills. This capability model can be used as a tool for gap analysis at the beginning of the GQM+ Strategies adoption process and in later phases of the adoption process to assess the progress of the adoption itself and to plan organizational training programs.

In fact, any new technology has to deal with technology adoption issues, and the GQM+ Strategies is not an exception. Therefore, it is important to study those issues and improve the approach accordingly.

9.3 Future Work

The work presented here can be further evolved using both empirical and analytical paradigms. Empirical research would involve studies of the applications of the GQM+ Strategies approach and its value-based extensions within organizations. Such empirical studies would allow for the evolution of the approach and the ability to measure its effects. It would be particularly interesting to have longitudinal studies with organizations that are using the GQM+ Strategies approach and its extensions. Analytical, new research would use the theories developed in this dissertation to further develop new theories and tools relevant to the GQM+ Strategies approach. In what follows, some research topics for the future work research directions are discussed.

The first group of topic deals with the limitations to the solution and to its application domain. This includes the efficient handling of complex grids, developing cost-benefit models for the GQM+ Strategies adoption, studying different adoption models with respect to socio-organizational factors, and the development of a GQM+ Strategies capability model. Investigation of the impact of the grid complexity on the grid development and maintenance processes, should look at factors that are affected by the grid complexity, and then hypothesize a model that describes their behavior and attempt to test that model empirically. Also, research can be conducted in a connection with previous research on the cost-benefit models for adopting the GQM+ Strategies.
approach. The third topic should investigate the human and socio-organizational aspects of adopting the GQM+ Strategies approach. More specifically, which organizational cultural models can positively impact the adoption of the approach? This study would help develop the adoption strategies for different organizations. Another topic of interest involves developing a GQM+ Strategies capability model. The model would specify various individual skills and organizational assets that are necessary in different phases of the GQM+ Strategies adoption process. These models could be used to perform gap analysis before the adoption, allowing organizations to plan accordingly.

The second group of topics addresses further development of the solution, and this includes evolving the GQM+ Strategies interpretation model, developing grid quality assurance procedures, and developing grid views for different roles. The GQM+ Strategies interpretation model is based on if-then rules, i.e. rule-based reasoning about the achievement of goals and strategies’ effectiveness. The rule-based reasoning is extremely effective in cases with the high level of determinism. However, when uncertainty dominates, a rule structure needs to be changed often as the new strategies and goals are added in a grid over time. A significant effort in updating the rule-structure can cause inconsistencies in the interpretation model, i.e. some logical errors. The formal model developed in this dissertation can be further evolved and used to enhance the interpretation model by substituting the rule-based reasoning with the causal probabilistic reasoning, which is highly effective for dealing with uncertainties. Using the formal model would also allow the development of algorithm-like procedures for checking the various quality attributes of the grid. For example, the grid should satisfy certain structural rules or logical constraints. During the GQM+ Strategies pilots it was noticed that different roles (e.g. managers and quality assurance officers) expected to see different levels of grid details. Therefore, different views of the grid should be created depending on the role of viewer. The solutions developed here can help with the presentation of different forms of information using different charts, diagrams, or even, colored graphs. This topic covers a whole area of visualizing and presenting the GQM+ Strategies grids.
References

9. Basili V (1994) GQM approach has evolved to include models.


192. van Solingen R & Berghout E (1997) Improvement by goal-oriented measurement: Bringing the goal/question/metric approach up to level 5. The European Software Engineering Process Group conference (E-SEPG), Amsterdam, The Netherlands, 1–11.
York, NY, USA, 1–12.


**Glossary**

\[ P(\sigma^i_j \mid x_J) \]

probability of having a successful strategy for some realization \( x_J \) of variables \( X_J \).

\[ \Xi_G \]

a mathematically formal structure that represents the GQM+Strategies grid defined over organizational goals \( G \).

\[ X_G \]

a set of all random variables assigned to organizational goals \( G \).

\[ M_G \]

the GQM based measurement schemes attached to the organizational goals \( G \).

\( i \)-strategies, \( \sigma(G)_j \)

strategies that address a goal \( G_i \) and refine into a collection of goals \( G_J \) at the next lower level.

\( i \)-superstrategy, \( \Sigma^i \)

a superstrategy represents the compound effect of all strategies that address a goal \( G_i \). An \( i \)-superstrategy specifies a complete set of goals that are refined during the grid derivation process with all defined strategies for a goal \( G_i \).

\[ G := \{ G_i \mid i \in N \}, N = \{0, 1, 2, ..., n\} \]

Set of all organizational goals.

\[ U_G \]

a set of all organizational units of goals \( G \).

**acceptable risk exposure, ARE**

the amount of risk that business owners (investors) are willing to take in order to materialize perceived benefits in a certain period of time.

**c-function** \( c(\cdot), V = u(G) \)

returns a set of all relevant contextual variables \( V \) of a goal \( G \).

**Chain Rule**, \( P(x_N) = \prod_{i=0}^{n} P(x_i \mid x_{i+1}, x_{i+2}, ..., x_n) \)

where \( x_N = \{x_0, x_1, ..., x_n\} \in S_N \), known as the chain rule of probability calculus, defines a way to calculate a probability of events occurring over the conditional probabilities of the events.
Conditional Probability, \( P(X_I = x_I | X_J = x_J) \)
the probability of occurring event \( \{X_I = x_I\} \) if an event \( \{X_J = x_J\} \) occurred, i.e. is given.

**Event** \( \{X_I = \bar{x}_I\} \)
denotes "\( G_i \) is not achieved".

**Event** \( \{X_I = x^*_I\} \)
denotes "\( G_i \) is achieved".

**Probability distribution for a goal**, \( P_{G_i}(s) \)
returns the probability (belief) that an organizational unit will achieve a goal \( G \) in a timeframe specified by the goal \( G \).

**Probability distribution marker**, \( \mu \)
empirically assessed probability by an expert.

**Probability distribution**, \( P \)
\[
P(x_N) = P(X_N = x_N) = P(\{X_1 = x_1, X_2 = x_2, ..., X_N = x_N\}) \quad \text{where} \quad N = \{0, 1, 2, ..., n\},
\]
and \( X_N = \{X_1, X_2, ..., X_n\} \), is the probability of occurring event \( \{X_N = x_N\} \).

**Relevant time moments**, \( T_G \)
is a set of all relevant time moments for goals \( G \).

**State space of a goal**, \( S_G \)
is a set of all possible states for a goal \( G \).

**t-function**, \( t(\cdot), T = t(G) \)
returns a timeframe of a goal \( G \).

**u-function**, \( u(\cdot), U = u(G) \)
returns organization unit \( U \) of a goal \( G \).

**Universe of discourse**, \( \mathcal{J}_G \)
represents all possible strategies and goals for an organization and its units and associated potential events that evaluate strategy effectiveness and goal achievement.
risk exposure, 84

stakeholders, 146
strategy assessment cards, 230

value, 145
organizational value management, 145
organizational value proposition, 147
value proposition, 146

VBSE, 27, 72, 144
4+1 theory, 76
Benefits Realization Analysis, 30
Business Case Analysis, 30
Change as Opportunity, 31
Concurrent System and Software Engineering, 31
Continuous Risk and Opportunity Management, 31
Stakeholder Value Proposition Elicitation and Reconciliation, 30
Value-Based Monitoring and Control, 31
Appendix 1 Interview Guide: Empirical Investigations of the GQM*Strategies Approach

1.1 Introduction

The main focus of our interview today is to understand, from your perspective, how the application of the GQM* Strategies method has affected your organization. We appreciate your feedback, and we will use it for the further method improvements. We understand that GQM* Strategies grid developed for your company is confidential; therefore, we are using this interview as an instrument of data collection that will make it possible to publish a valid study without disclosing company confidential data. The interview will be recorded and transcribed (as stated in the interview consent form). At any point during the interview you can request to go off the record (to stop recording the interview). After the interview, the confidentiality of the material will be dealt with the following protocol. The interview summary will be sent to interviewee for a review and checking on confidentiality. All information will be treated as confidential and published in an aggregated and anonymous form, with a previous interviewee acceptance before publishing.

We will try to limit the interview time to 1 hour. During the interview we will go through a set of context related questions, questions concerning your role/work in the company, and questions regarding your observations about the application of the GQM* Strategies method.

<table>
<thead>
<tr>
<th>NOTE FOR INTERVIEWER</th>
</tr>
</thead>
<tbody>
<tr>
<td>During the interview we will go through following types of questions:</td>
</tr>
<tr>
<td>– Warm-up questions: demographic and context setting questions</td>
</tr>
<tr>
<td>– General questions: describing your work responsibilities, typical problems that you are dealing with, and the situation that preceded the GQM+Strategies work</td>
</tr>
<tr>
<td>– Observation questions: your observation regarding the application of the method to date. What you found useful, what was good, and what was not.</td>
</tr>
<tr>
<td>– Follow-up questions: more in detail elaboration of the method’s effects</td>
</tr>
<tr>
<td>– Wrap-up questions: closing questions and final question.</td>
</tr>
</tbody>
</table>

Any questions before we begin?
1.2 Interview Questions

NOTE FOR INTERVIEWER
The format of a question:
Q#.#: Question to be asked by interviewer. [Clarifications for interviewer]
✓ Sub-questions, topics to be covered, focusing questions. [Clarifications for interviewer]

1.2.1 Warm-Up Questions

Q1.1: For which COMPANY are you currently working? [It is for the record; if interviewer is familiar with the company he/she can start answering the question and ask for confirmations.]
✓ In brief: industry type and domain (market)?
✓ Main products or services?
✓ Size of the company?

Q1.2: Tell us something about your work experience?
✓ How long have you been with the company?
✓ What other company experience have you had? Same industry and domain? [Relevant for the interviewee’s opinion about applicability of the method.]
✓ What is your current position/role in the company? [Just name the role, do not go into role description details.]
✓ How long have you been in your current position/role?
✓ Have you had different positions/roles before? [Different roles can perceive different problems/benefits of the method.]
✓ Measurement exposure, have you been using metrics?

Q1.3: Tell us something about your professional background? [The focus is on the professional background; is it managerial, or engineering, or both.]
✓ Professional degrees.
✓ Specific certificates.
1.2.2 General Questions

Q2.1: For what kinds of decisions are you responsible and what kind of support tools are you using for decision making? [The focus is on the aspects of the work that are relevant for the GQM Strategies application.]

✓ Are you using any decision support systems?
✓ Do you always (when you make a decision) understand the organizational scope and implications of the decision?
✓ While performing your daily work tasks, what are typical problems/issues/challenges that you have to deal with? [Limit answer on major issues, e.g. top two issues, and preferably focus on/select issues that are relevant to the GQM+Strategies application.]

Q2.2: What methods or approaches (excluding GQM Strategies) are used in your company to define and communicate business goals and strategies? [The focus is on existing methods that are alternatives to GQM Strategies.]

✓ How are you informed about business goals and strategies? [If interviewee cannot name approaches and methods]
✓ How are business and operational goals controlled and monitored?
✓ Do you see some issues with the current methods and approaches?

Q2.3: What is a common practice in your company for collecting quantitative data from different projects and units? [It might happen that there are no defined practices, but still some data are probably generated somewhere (e.g., financial indicators, time sheets, bug reports, etc.).]

✓ What kind of metrics (e.g., financial indicators, product metrics, etc.) are you using? [Kinds of metrics can identify groups (stakeholders) who have interests in measurements. Identify those groups for a later reference (Q4.2).]
✓ Do you see some issues with the current practices? [Identify issues for a later reference.]

Q2.4: What were the main objectives in trying GQM Strategies? [Identified issues from previous answers]

✓ What were the main motivators for trying GQM+Strategies? [Other motivators, except the identified issues.]
Q2.5: Can you summarize your involvement with GQM+Strategies work (a grid building activities)? [Interviewer should start summarizing joint activities or sessions (e.g., tutorials, workshops)]

✓ What were your activities in between the joint sessions?
✓ Did you do group work (e.g. brainstorming)?
✓ Did you do individual work?

During our joint collaboration regarding the GQM+Strategies method application in your company (as you elaborated in the previous answer, Q2.5) you have participated in the development of a pilot GQM+Strategies grid (from now on, we will refer it as the grid).

1.2.3 Observation Questions

Q3.1: Was the process of building the grid helpful for your organization? [The focus is on the process (activities, set of tools (e.g., templates), and additional materials).]

✓ What activities of building the grid did you find most useful?
✓ Why are those activities useful?
✓ What problems did you have with the grid building process?

Q3.2: Is the resulting grid helpful for your organization? [The focus is on the grid as an artifact and the utility of that artifact.]

✓ What features of the grid did you find particularly useful?
✓ What problems were you able to solve with the grid? Why? [Note down the most interesting issues and problems for a later reference.]

Q3.3: Would you be able to evolve the resulting grid (e.g. add goals and strategies) on your own? [The focus is on the abilities and skills to perform the grid derivation steps.]

✓ Where would you require more help?
✓ (If not able) What are the main problems?
✓ Reflect answers in tutorial and grid evaluation questionnaires? [If interviewee has answered them]
1.2.4 Follow-Up Questions

Q4.1: How did the process of building grid or the resulting grid help you to address the specific problems you mentioned?

✓ Prompts: identified issues in question Q3.1 and Q3.2.

Q4.2: whom in the company do you think the grid will be of value? [Consider various groups and roles in the organization.]

✓ How and why?
✓ Is it useful for your work? How and why?

Q4.3: How has building the grid and the resulting grid impacted your way of thinking about your business and daily work? [EU understanding of the term "business" is strategic or corporate level thing, while in US business also includes operations and daily routines.]

✓ (If no impact then) Why?

Q4.4: Do you think that the grid can support higher-level decision-making by providing relevant project-level information in your company? [The focus is on decision making; if an interviewee is a manager question addresses the utility of information provided by a grid; and, if the interviewee is not a manager, the question addresses the ability of the grid to capture relevant information.]

✓ How?

Q4.5: How did the grid building process and the resulting grid help you communicate goals and strategies across your company? [The focus is on communication between different organizational levels, including transparency and alignment.]

✓ How well goals and strategies suit your environment?
✓ Does the grid support you in getting a clearer picture of the goals and strategies of your organization? [Transparency]
✓ Does the grid represent a logically consistent set of goals and measurement data? [Alignment]
✓ How can the grid help you in planning and implementing goals and strategies for your organization or project?
✓ How well do you think the grid captured information about your internal and external (i.e. industry) context? [Transparency]

Q4.6: How well did the grid help you to quantify organization’s goals and strategies? [The focus is on measurability.]
✓ To identify unquantifiable goals or risky goals.
✓ To interpret success/failure of goals and strategies.
✓ How well did the grid help you to understand the feasibility of goals and strategies?

1.2.5 Wrap-Up Questions

Q5.1: Is GQM\(^*\) Strategies a valuable addition to your organizational repository of managerial methods and approaches?
✓ Comparing existing methods used by organization (Q2.3) and GQM\(^*\) Strategies; what are advantages of the GQM\(^*\) Strategies?
✓ What are disadvantages?

Q5.2: Has anyone else in your organization expressed an interest in GQM\(^*\) Strategies?
✓ Do you think it will be used in other parts of your organization?
✓ (If not) What are reasons?

Q5.3: Do you think GQM\(^*\) Strategies would be helpful to other organizations in your domain?
✓ (If not) What are reasons?

Q5.4: Any related issues that we missed but you would like to reflect on?
✓ Are those minor concerns or serious threats?

NOTE FOR INTERVIEWER
Stop the interview recording.

Thank you for your time and participation in this interview. After we analyze the interview results, we will send you a summary of findings for checking on any inconsistencies with today’s interview discussion. We hope you can get us feedback within two weeks after we send you the summary of findings. Also, if you would like we can provide you a full transcript of the interview.
Appendix 2 BigCorp: A Comprehensive GQM*Strategies Example

This example is based on real-world situations encountered during the GQM*Strategies applications with ICT industry. Meaning that the parts of the grid contain strategies for addressing some specific problems, and those problems were encountered during the GQM*Strategies applications in different organizations. Here, the exemplar grid presents them like they are all addressed by a single organization or an imaginary company that is referred as BigCorp. The goal of this example is to illustrate the steps of the GQM*Strategies process and feasibility of the proposed method extensions.

The example follows the structure of the GQM*Strategies process (section 2.3.2), that is defined by QIP phases of initialize, characterize, set goals, choose process, execute process, analyze results, and package and improve.

BigCorp is a large international company in the business of embedded systems. Company is successful in producing mobile software applications for the mobile hardware platform that the company developed and owns. The applications are developed mainly for the defense industry. The main advantage of the BigCorp is a highly robust and durable hardware that can be used in industrial environments or in the field (e.g. mining operations, military purpose). BigCorp offers a wide range of COTS\textsuperscript{45} for their clients, but it also offers a custom software development for their hardware platform.

2.1 Initialize

During the initialize phase the responsibilities of the involved people were clarified and a draft schedule was created. The schedule includes:

– **Method tutorial** for the key persons and **method champion**.
– Method presentation for all people that will be involved with the grid development.
– **Five workshops** for developing grid with different groups of people (i.e. units)
– Off-line work, between workshops, involving method champion for defining measurement schemes.
– Feedback session for presenting the whole grid

\textsuperscript{45}Component Of-The-Shelf
The first meeting with BigCorp representatives included a short presentation of the approach.

2.2 Characterize

In order to determine the scope of the application, the organizational structure of the BigCorp is used (Figure 38).

Fig 38. BigCorp organizational structure.

The BigCorp’s management board involves CEO and four business units’ directors. The BigCorp has four units: marketing & PR, SoM (Software on Mobile), MHP (Mobile Hardware Platform), and customer support unit. The director of the SoM (Software on Mobile) business unit initiated the adoption of the GQM+Strategies, because he felt that his R&D unit is capable of delivering great results, but those capabilities and results are not visible at the top level. The scope of the application involves eight units and management board (marked as $U_0$). All involved units delegated persons who will be participating the grid derivation process. The representatives have the good understanding of their unit’s goals and operations. They act as goal owners in the front
of their units and groups. The R&D manager volunteered to take the role of the method champion.

**Notation Remarks** Throughout the example, the elements of BigCorp grid are enumerated in following way. Units are marked with "U" indexed with unit number, e.g. $U_0$, $U_1$, etc. Organizational goals are marked with indexed capital letter "G," e.g. $G_0$, $G_1$, etc. The context and assumption elements are marked with indexed capital letters "C" and "A," e.g. A1, C2, A2, etc. Strategies are marked with Greek symbol sigma "$\sigma^{u\ldots\,m}_{n}$" where the upper-script number designates the upper-level goal, and under-script numbers designate the refined goals by the strategy. For example, strategy $\sigma^{2}_{5,6}$ addresses goal $G_2$ and refines into goals $G_5$ and $G_6$.

### 2.3 Set Goals

Organizational goal hierarchy was developed using several workshops, i.e. group sessions. Where, each session had two parts. First, brainstorming and eliciting potential goals and strategies. Second, the selected goals and strategies were formalized using the GQM+Strategies templates.

The overview of the goals is given in Figure 39, and a complete list of context and assumption elements is given in Tables 26 and 27.

#### 2.3.1 Workshop with Top-level Management

The management board felt that BigCorp business is mature enough, and that it can be expanded. Therefore, they set the top level goal as increasing net earnings ($G_0$), the goal is formalized with the GQM+Strategies template (Table 13). BigCorp serves very specific industries, e.g. mining and military solutions (C1), which are not reacting in a same way to the global crises as the consumers market (C2). Thus, there is an opportunity to gain new customers by providing BigCorp products that can substitute some existing and more expensive products (A1). Therefore, a strategy $\sigma^{0}_{1}$ to increase sales, i.e. sale more to the existing clients or to win new clients was selected.

Furthermore, the management board identified some process inconsistencies that cause project delays and additional rework (C3) to be a potential risk; especially, if the company is going to serve several times more clients than current. Therefore, another
strategy $\sigma_2^0$ to improve production processes was also selected. Improving process performance should reduce costs, and increase number of billable hours (A2).

The strategy of increasing sales refines goal $G_0$ into $G_1$: *increase sales in all regions*. While, the second strategy refines the top-level goal into $G_2$: *improve process performance*. The assumption (A3) is that increasing sales by 20% CAGR\(^{46}\) ($G_1$) and reducing 5% costs on manufacturing and 15% costs on software development ($G_2$) would result with planned increase of net earnings ($G_0$).

### 2.3.2 Workshop with Marketing & PR unit

The marketing representatives agreed that the most important goal is to *increase market awareness about BigCorp brand and products* ($G_3$). The marketing unit has established marketing channels that can reach all potential clients (C4), and the contacts with potential large clients already exist (C5). The assumption (A4) states that usual marketing tools (e.g., promotions, specialized events, exhibitions, seminars) are sufficient to achieve $G_3$.

### 2.3.3 Workshop with SoM unit

The SoM unit consists of several subunits that are involved with the grid derivation. The sales unit is a part of the SoM because the end product are software products and services, i.e. the hardware platform is never sold without software applications. The sales unit will focus on executing sales goal $G_1$ by redefining BigCorp’s product portfolio (those strategies are not included in this example).

The development of the new software and customization of the existing software applications is done by R&D unit. Applications developed by BigCorp are limited in complexity (C6), and they are developed as a part of small scale projects, i.e. there is no need for "heavy" software development processes (A5). Therefore, the strategy $\sigma_4^2$ to adopt agile practices was selected. The strategy refined goal $G_2$ into $G_4$: *establish agile SW development process*. Assumptions are that current rework can be reduced by introducing short iterations (A6), and that software quality can be further improved by implementing test-driven practices (A7).

---

\(^{46}\)Compound Annual Growth Rate

274
### BigCorp Goals

<table>
<thead>
<tr>
<th>Goal</th>
<th>Activity</th>
<th>Focus</th>
<th>Object</th>
<th>Magnitude</th>
<th>Timeframe</th>
<th>Scope</th>
<th>Constraints</th>
<th>Relations</th>
<th>Org. Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>G0</td>
<td>Increase</td>
<td>Net Earnings</td>
<td>All product lines</td>
<td>160%</td>
<td>5 years</td>
<td>All units</td>
<td>Global crisis</td>
<td>-/-</td>
<td>U0</td>
</tr>
<tr>
<td>G1</td>
<td>Increase</td>
<td>Sales</td>
<td>All regions</td>
<td>20% CAGR</td>
<td>5 years</td>
<td>All units</td>
<td>Global crisis</td>
<td>-/-</td>
<td>U5</td>
</tr>
<tr>
<td>G2</td>
<td>Improve</td>
<td>Process performance</td>
<td>production processes</td>
<td>-5% on manuf. cost -15% on SW develop. cost</td>
<td>3 years</td>
<td>SoM, MHP units</td>
<td>Global crisis</td>
<td>-/-</td>
<td>U0</td>
</tr>
<tr>
<td>G3</td>
<td>Increase</td>
<td>BigCorp brand and products</td>
<td>Market Awareness</td>
<td>70% of big potential clients would recognize BigCorp brand</td>
<td>2 years</td>
<td>Marketing and PR unit</td>
<td>-/-</td>
<td>-/-</td>
<td>U1</td>
</tr>
<tr>
<td>G4</td>
<td>Establish</td>
<td>Agile SW Process Compliance</td>
<td>Develop. process</td>
<td>less than 10% of rule violation for 95% process coverage</td>
<td>2 years</td>
<td>SoM unit</td>
<td>-/-</td>
<td>-/-</td>
<td>U2</td>
</tr>
<tr>
<td>G5</td>
<td>Improve</td>
<td>Efficiency &amp; effectiveness</td>
<td>Manufact. process</td>
<td>15% of cost factor reduction</td>
<td>2 years</td>
<td>MHP unit</td>
<td>-/-</td>
<td>-/-</td>
<td>U3</td>
</tr>
<tr>
<td>G6</td>
<td>Establish</td>
<td>Collective code ownership</td>
<td>Product X</td>
<td>Min truck factor at 70% of code coverage larger than 10</td>
<td>1 years</td>
<td>R&amp;D unit</td>
<td>-/-</td>
<td>-/-</td>
<td>U6</td>
</tr>
<tr>
<td>G7</td>
<td>Reduce</td>
<td>Process variation</td>
<td>Manufact. proc P1,P2</td>
<td>0.3sigma less variation</td>
<td>1.5 years</td>
<td>Manufactur unit</td>
<td>-/-</td>
<td>-/-</td>
<td>U7</td>
</tr>
<tr>
<td>G8</td>
<td>Improve</td>
<td>Quality</td>
<td>Customer interaction process</td>
<td>20% less complaints</td>
<td>2 years</td>
<td>Customer support unit</td>
<td>maintaining current operational budget</td>
<td>-/-</td>
<td>U4</td>
</tr>
</tbody>
</table>
are necessary, therefore the strategy $\sigma_{6,7,8}^4$ is to establish key agile practices. This strategy leads to the three lower-level goals, $G_6$: *establish collective code ownership*, and two other goals $G_7$ and $G_8$ that deal with the other two key agile practices.

### 2.3.4 Workshop with MHP unit

The MHP unit identified the strategy $\sigma_2^5$ to reducing manufacturing costs to be the most effective for the upper-level goal. An assumption is that manufacturing process can be further optimized to gain sufficient cost reduction (A9). The strategy leads to the unit’s goal $G_5$: *improve efficiency and effectiveness of the manufacturing process*. It was assumed that manufacturing process can be fine tuned and further optimizes (A10). At manufacturing level the statistical process control (SPC) techniques are used and baseline data for key manufacturing processes exist (C7). Therefore, strategy $\sigma_9^5$ to decrease process variation was selected, and resulted in the lower-level goal $G_9$: *reduce process variation for 0.3σ*. Assumption is that reducing process variation for selected sub-processes will result with 5% cost factor reduction (A10).

### 2.3.5 Workshop with Customer Support unit

The customer support unit is important for maintaining the current pool of customers. Some deficiencies were identified that are related to the customer interaction. The responsiveness to the customer requests is not satisfactory (C8), and assumption is that due to inaccurate information customer support runs additional information checks (A11). Therefore, strategy $\sigma_{10}^1$ to increase responsiveness of the customer support refines goal $G_1$ into goal $G_{10}$: *improve quality of customer interaction process*. An assumption (A12) is that improving quality of the customer support is needed to accommodate new sales to existing customers.

### 2.3.6 Defining GQM measurement schemas

In Figure 39 is given a list of BigCorp goals formalized with the GQM+ Strategies template. For the each goal a measurement scheme, i.e. GQM Graph, is defined. The process of defining graphs was done mainly off-line, after the workshops. The method champion assisted with collecting available data and metrics.
Table 26. Context and assumptions for the BigCorp grid.

<table>
<thead>
<tr>
<th>C/A#</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>BigCorp provides industry and military mobile solutions based on their highly robust mobile hardware platform.</td>
</tr>
<tr>
<td>C2</td>
<td>The global crises is affecting less the military industry.</td>
</tr>
<tr>
<td>A1</td>
<td>BigCorp products can substitute some existing and more expensive products.</td>
</tr>
<tr>
<td>C3</td>
<td>Existing process inconsistencies are causing project delays and rework.</td>
</tr>
<tr>
<td>A2</td>
<td>Improving process performance can increase the number of billable hours and reduce production costs.</td>
</tr>
<tr>
<td>A3</td>
<td>Increasing sales for 20% CAGR and reducing costs of production (5% from manufacturing and 15% from software development) is sufficient to increase net earnings for 160% in five years.</td>
</tr>
<tr>
<td>C4</td>
<td>The marketing unit has established marketing channels that can reach all potential clients.</td>
</tr>
<tr>
<td>C5</td>
<td>Contacts with potentials “big” clients already exist.</td>
</tr>
<tr>
<td>A4</td>
<td>Marketing department has all needed skills and marketing tools.</td>
</tr>
<tr>
<td>C6</td>
<td>Mobile hardware platform does not allow development of complex software applications.</td>
</tr>
<tr>
<td>A5</td>
<td>There is no need for a heavy project administration for developing software components.</td>
</tr>
<tr>
<td>A6</td>
<td>Introducing short project iterations will significantly reduce rework.</td>
</tr>
<tr>
<td>A7</td>
<td>Test-driven development practices can provide sufficient level of quality.</td>
</tr>
<tr>
<td>A8</td>
<td>Key agile practices for R&amp;D are: test-driven development, continuous refactoring, and collective code ownership.</td>
</tr>
<tr>
<td>A9</td>
<td>Optimizing manufacturing processes can gain significant cost reductions.</td>
</tr>
<tr>
<td>A10</td>
<td>Fine tuning (i.e. no restructuring) manufacturing processes can provide 5% cost reductions.</td>
</tr>
<tr>
<td>C7</td>
<td>Manufacturing processes are managed using the SPC techniques, and baseline data exist for the key manufacturing processes.</td>
</tr>
<tr>
<td>C8</td>
<td>Responsiveness to the customer request is not satisfactory, many complains from customers.</td>
</tr>
<tr>
<td>A11</td>
<td>Due to inaccurate information customer support runs additional information checking that cause response delays.</td>
</tr>
<tr>
<td>A12</td>
<td>The quality of the customer support is the main obstacle for selling more to existing clients.</td>
</tr>
</tbody>
</table>

For the each organizational goal a measurement goal is defined using the GQM template. For example, the top-level goal is measured with following measurement goal:

\[ M_{G0} : \text{Analyze all product lines in order to evaluate with respect to } 160\% \text{ increase of net earnings in the period of 5 years} \text{ from the point of view of BigCorp.} \]

Where context is documented with context and assumption elements in Table 26.

The measurement goal \( M_{G0} \) leads to the following questions:
Fig 40. BigCorp the GQM+Strategies goal hierarchy.

Q1: What were net earnings for the past year?
Q2: What are current net earnings?

The question Q1 sets the baseline for measuring the increase of the net earnings. If the net earnings are greater than 160% of the baseline than the organizational goal $G_0$ is achieved.

All measurement goals comprise the GQM graph associated with BigCorp’s goal hierarchy. Measurement goals are given in Figure 41 and Figure 42.

In parallel with defining GQM goals the interpretation models is constructed.

//Example of the GQM+Strategies interpretation model
interpretation_model_start:
IF \( \frac{CNE(y)}{BNE} \geq 1.6 \) THEN // Comment
Goal \((G_0)\) of 160% net earnings is ACHIEVED;
ELSE IF \( \sum_r \frac{BR(r,y)}{BS} \geq 1.2\ AND\ CrMNF \geq 0.05\ AND\ CrSW \geq 0.15 \) THEN //Comment
Goal \((G_1)\) of 120% CAGR sales increase is ACHIEVED;
Goal \((G_2)\) of improving process performance is ACHIEVED;
Check assumptions connecting \(G_0\) with \(G_1\) and \(G_2\);
ELSE IF lower-level goals achieved THEN
re-evaluate level-i decisions
Check assumptions connecting \(G_1\) and \(G_2\) with lower-level goals;
ELSE
Reconsider strategies (i.e. action plans) for lower-level goals;
ELSE
Check context and assumptions for lower-level goals;
<table>
<thead>
<tr>
<th>Org. Goal</th>
<th>Measure. Goal</th>
<th>Object</th>
<th>Purpose</th>
<th>Quality</th>
<th>Aspect</th>
<th>Viewpoint</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>G0</td>
<td>MG0</td>
<td>All product lines</td>
<td>Evaluate</td>
<td>Net Earnings increase of 160% in 5 years</td>
<td>BigCorp</td>
<td>See context and assumptions</td>
<td></td>
</tr>
</tbody>
</table>

Q1: What were net earnings for the past year (baseline)? Q2: What are current net earnings?

M0.1: \( CNE(y) \) = corporate net earnings for the year \( y \) in euros.
M0.2: \( BNE = CNE(2010) \), baseline net earnings

| G1 | MG1 | All regions | Evaluate | Sales increase of the 20% CAGR in 5 years | BigCorp | See context and assumptions |

Q1: What are the latest sales figures by regions (baseline)? Q2: What are current sales figures?

M1.1: \( SBR(r, y) \) = total sales in the year \( y \) in euros for the region \( r \).
M1.2: \( BS = \sum SBR(r, 2010) \) for all \( r \) (baseline sales for all regions)

| G3 | MG3 | Market awareness | Evaluate | BigCorp brand and product visibility; 70% big clients to be aware of the BigCorp brand | BigCorp | See context and assumptions |

Q1: Who are the major clients (i.e. "big" clients)? Q2: How many of them received BigCorp product tutorial?

M3.1: \( BAW = \# \) of "big" clients / \# given tutorials %

| G3 | MG3 | Market awareness | Evaluate | 70% big clients to be aware of the BigCorp brand | BigCorp | See context and assumptions |

Q1: Who are the major clients (i.e. "big" clients)? Q2: How many of them received BigCorp product tutorial?

M3.1: \( BAW = \# \) of "big" clients / \# given tutorials %

| G4 | MG4 | SW development process | Evaluate | Agile SW process compliance | SoM unit, QA team | See context and assumptions |

Q1: What is the process conformance coverage (per iteration or project)? Q2: How many violations per iteration or project? Q3: Is process well defined?

M4.1: \( RCV = \# \) selected rules / Total # process rules (templates) %
M4.2: \( RVI = \# \) violated rules / Total # checked rules %
M4.3: \( RCH = \# \) rule changes (versions) per rule
M4.4: \( GoR(\ell) = RCH(\ell) \) per iteration or project

| G10 | MG10 | Customer interaction process | Evaluate | Improve quality by having 20% less complaints in 2 years | Customer support unit | See context and assumptions |

Q1: How many customer complaints exist regarding the interaction process?

M10.1: \( CCI(T) = \# \) customer complaints about the interaction process in time span \( T \).
<table>
<thead>
<tr>
<th>Org. Goal</th>
<th>Measure. Goal</th>
<th>Object</th>
<th>Purpose</th>
<th>Quality Aspect</th>
<th>Viewpoint</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>G5</td>
<td>M5</td>
<td>Manufacturing processes</td>
<td>Evaluate efficiency (i.e., degree to which outputs are produced) and effectiveness (i.e., specification tolerance)</td>
<td>MQHP unit, QA team</td>
<td>See context and assumptions</td>
<td></td>
</tr>
</tbody>
</table>

**Questions / Metrics**

Q1: What is the efficiency of the process?
Q2: What is the effectiveness of the process?

M5.1: \( \text{VOP}(x) = \text{voice of the process x} \)
M5.2: \( \text{VOC}(x) = \text{voice of the customer for process x} \)
M5.3: \( \text{ST}(x) = Z_u(x) + Z_l(x) \) (specification tolerance in sigma units)

<table>
<thead>
<tr>
<th>Org. Goal</th>
<th>Measure. Goal</th>
<th>Object</th>
<th>Purpose</th>
<th>Quality Aspect</th>
<th>Viewpoint</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>G6</td>
<td>M6</td>
<td>All products</td>
<td>Evaluate</td>
<td>Collective code ownership</td>
<td>R&amp;D unit, Dev. teams</td>
<td>See context and assumptions</td>
</tr>
</tbody>
</table>

**Questions / Metrics**

Q1: How many members of project team are working per iteration? Q2: How many members of the team in the current iteration worked in the previous iteration? Q3: How many new members of team joined the team? Q4: What is the level of teams vulnerability to people turn over?

M6.1: \# project members per iteration
M6.2: \# current members that worked on previous iteration / total \# current members %
M6.3: \# new members / total \# current members %
M6.4: \( tf(\text{min}, 70\%) \): truck factor metric

<table>
<thead>
<tr>
<th>Org. Goal</th>
<th>Measure. Goal</th>
<th>Object</th>
<th>Purpose</th>
<th>Quality Aspect</th>
<th>Viewpoint</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>G9</td>
<td>M9</td>
<td>Manufacturing processes P1, P2, P3</td>
<td>Control Process variation</td>
<td>MHP unit, QA team</td>
<td>See context and assumptions</td>
<td></td>
</tr>
</tbody>
</table>

**Questions / Metrics**

Q1: Is the process performing within control limits? Q2: Is process showing signs of instability?

M9.1: \( \text{UCL} = \text{upper control limits in sigma units} \)
M9.2: \( \text{LCL} = \text{lower control limits in sigma units} \)
M9.3: \( \text{CL} = \text{center line} \)
M9.4: Classification of data patterns (P1...Pn)
The statement "re-evaluate level-i decisions" is often used in interpretation model, and it means to check context and assumptions (which can be seen as rationales for decisions) if there are any changes or assumptions that proved to be wrong. The new/changed set of assumptions most probably will result with new level-i decisions.

2.4 Choose Process

In order to make plans for implementing organizational goals and strategies, the cost–benefit analysis was performed in order to estimate costs and benefits associated with organizational strategies. In addition, risks associated with organizational goals were evaluated. The results of analysis were used to plan budget for the grid execution.

2.4.1 Cost–Benefit Analysis

The cost–benefit analysis was done by the people who developed the grid, i.e. they know the best what kind of costs are involved with implementing goals, and what kind of benefits can be expected.

Figure 43 illustrate the results of the cost–benefit analysis, showing the cost benefit estimates for the each goal. The cost and benefits that are inherited from lower level goals are presented by numbers on the connecting edges. The analysis is done for the period of time indicated by top-level goal. In the case when the analysis period is longer than a goal’s timeframe, e.g. goal \( G_5 \) has timeframe of 2 years and time period of analysis is 5 years, then the goal’s maintenance has to be considered and included in total estimates for the goal together with the goal specific and inherited costs and benefits. For example, the goal \( G_5 \) will be maintained for 3 years after it is achieved.

The results of analysis are given in absolute numbers (e.g., EU currency), and in order to produce absolute numbers the whole case was scaled for following parameters: 100 people, 120 hours per employee/month, internal cost of employee/hour is 10 EUR. The cost–benefit analysis resulted with additional context and assumption elements, as documented in Table 27. Context and assumptions identified during the cost–benefit analysis are differentiated with an asterisk sign, e.g. A3* or C2*, from remaining context/assumption elements.

The development and QA team (R&D unit) estimated that main costs for goal \( G_6 \) (establishing collective code ownership) will come from needed redundancy, e.g. paired programming, (A2*) and that will cost 336k \( \text{€} \) (where, \( 1k = 1000 \)) for five years. At the
same time, new practices will save hours spent on familiarizing new team members (A5*), which for the current people turnover is about 100k € in five years. While, for goals $G_7$ and $G_8$ the main costs are licensing additional software development tools (C1*). Test-driven practices, $G_8$, are expected to reduce defect slippage (A6*), using the current data for the effort spent on fixing defects it is estimated that savings could be 400k € for the period of five years. All those costs and benefits are aggregated by goal $G_4$ (establish agile SW process compliance), in addition to those estimates, the QA team planned resources for training and followup meetings (A4*). Therefore, total costs for goal $G_4$ are 446k € and expected benefits 500k €, all for a five-year period.

All teams performed the same analysis for their goals and strategies. In the case of goal $G_{10}$ it was estimated that costs and benefits are equally shared by upper-level goals $G_1$ and $G_2$, i.e. improving customer interaction process equally contribute to a better sales with existing clients and improves process performances; therefore costs and benefits are shared (A9*).
### Table 27. Extended set of BigCorp’s context and assumptions with cost–benefit analysis.

<table>
<thead>
<tr>
<th>C/A#</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1*</td>
<td>The marketing task force team estimated costs for their operation (e.g. meetings, internal seminars) as 50K for five years period.</td>
</tr>
<tr>
<td>A2*</td>
<td>Implementing collective code ownership, requires certain level of redundancy (i.e. several people to work on same problem) that will increase internal costs per hour for 3.5%.</td>
</tr>
<tr>
<td>C1*</td>
<td>Implementing test-driven development and continuous integration practices requires some third party tools.</td>
</tr>
<tr>
<td>C2*</td>
<td>Fine tuning the manufacturing processes requires calibration techniques that are expensive and provided by third party.</td>
</tr>
<tr>
<td>C3*</td>
<td>Fine tuning the manufacturing processes requires calibration techniques that are expensive and provided by third party.</td>
</tr>
<tr>
<td>A3*</td>
<td>The process control team estimated costs for their operations as 50K for five years.</td>
</tr>
<tr>
<td>A4*</td>
<td>The software QA team estimated their costs for tutoring the agile practices (i.e. meetings, presentations, courses, followup) to be 80K for five years.</td>
</tr>
<tr>
<td>C4*</td>
<td>Costs of management board meetings for five years (10 meetings, 2 per year) are 50K.</td>
</tr>
<tr>
<td>A5*</td>
<td>New practices for establishing collective code ownership can save 50% of familiarization costs for new employees.</td>
</tr>
<tr>
<td>A6*</td>
<td>Test-driven development practices can further reduce bug slippage for 10%.</td>
</tr>
<tr>
<td>A7*</td>
<td>New manufacturing schedule will increase the use of resources for 18%.</td>
</tr>
<tr>
<td>A8*</td>
<td>Internal cash flow generated from sales in first year 35%, second 40%, third through fifth 30%.</td>
</tr>
<tr>
<td>A9*</td>
<td>Goal $G_{10}$ equally contributes to the goals $G_1$ and $G_2$.</td>
</tr>
<tr>
<td>A10*</td>
<td>The targeted market will continue to be affected by the crisis.</td>
</tr>
<tr>
<td>A12*</td>
<td>Resistance to change is high</td>
</tr>
<tr>
<td>A13*</td>
<td>Data migration from old CS system to new one</td>
</tr>
</tbody>
</table>

The marketing team allocated 350k € for implementing goals and strategies associated with goal $G_3$ (not shown in the Figure 43). This includes exhibitions, promotions, and client’s visits. While costs for goal $G_3$ are 50k €, including team meetings and internal seminars (A1*).

The sales manager provided the model for converting expected sales (i.e. the $G_1$ magnitude) into the internal cash flow, i.e. the money that remains after some taxes and paying suppliers and distributors (A8*). According to that model, the expected benefits of the goal $G_1$ is 11050k € in five years (about 11 millions €).
The top level goal $G_0$ aggregates all costs and benefits for the five years of strategies implementation. Total costs are 1046k €, including costs of board meetings for that period (C4*); and, total benefits are 12320k €. The cost–benefit analysis provided BigCorp management board with the following information:

- Selected strategies and goals can realize about 85% of the initially targeted increase by the 160% of the net earnings. Therefore, the management board can correct the target, or choose alternative strategies if the increase is not sufficient with selected strategies.
- For selected strategies, and given size of the investment and expected benefits for a five-year period is for every invested euro should playback ten times, i.e.

$$ROI = \frac{12320 - 1046}{1046} = 10.8.$$ 

This represents a rough estimate of the ROI factor for BigCorp’s top-level goal and strategies. The effects of time on the value of money are not accounted.

### 2.4.2 Risk Analysis

The management board is responsible for making investment decisions; therefore, they specified an acceptable risk model.

The size of investment and benefits was quantified as: 1–small, 2–medium, and 3–large. The investment size ($I$) is categorized in respect to the absolute number of investment units (money or any equivalent), while the benefits size ($B(I)$) is categorized as the relative quantity of the invested amount. Given the available resources of €1 million, the company management board specified an investment of less than €0.5 million as small, an investment of greater than €0.5 million and less than €0.8 million as medium, and greater than €0.8 million as large. The common understanding is that small benefit is less than 30% of investment, medium is between 31% and 65%, and large benefit is if it is greater than 65% of investment.

In order to produce the acceptable risk matrix (Equation 37), the board members were asked to specify acceptable risk exposure for each pair of investment size and benefit size. Risk exposure was quantified using a five-point scale from VH (very high risk), most probable to occur with significant consequences, to VL (very low risk), unlikely to occur with no severe consequences. Therefore, the ARE matrix is:
\[
ARE = I \times B(I) = \begin{bmatrix}
M & L & VL \\
H & M & L \\
VH & H & M \\
\end{bmatrix},
\]

(37)

where the columns represent investment size and the rows the benefits size. Risk exposures: \(VH\) is very high, \(H\) is high, \(M\) is medium, \(L\) is low, and \(VL\) is very low.

The likelihood of an assumption (or assumption certainty) is quantified using a three-point scale: 1–low, 2–medium, and 3–high. The goal impact is also specified as: 1–low, 2–moderate, 3–significant. Based on that, risk exposure is calculated from the \(RE\) matrix (Equation 38). Risk exposure is quantified with the same five-point scale that was used for the acceptable risk exposure.

\[
RE(x) = L(y) \times I(y, x) = \begin{bmatrix}
VL & L & M \\
L & M & H \\
M & H & VH \\
\end{bmatrix},
\]

(38)

where the columns represent the likelihood of an assumption \(y\), and the rows are the impact of assumption \(y\) on contentment of business goal \(x\).

The key assumption for the top-level goal \(G_0\) is that BigCorp can substitute a more expensive product with their own in the market (A1). This is a positive assumption. Therefore, the experts were asked to assess the likelihood of failing to substitute the expensive products. The assessment was that the likelihood was low and the impact of the assumption on the goal’s success was moderate, i.e. \(RE(G_0) = low \times moderate = L\).

Table 28 gives the results of goals risk analysis. According to the results, only goals \(G_6\) and \(G_{10}\) are considered risky. The goal of establishing collective code ownership \(G_6\) is perceived as risky (\(RE(G_6) > ARE(G_6)\)) because it depends on the ability of people changing their habits (A12*), which was assessed as of high importance for the goal’s success (the likelihood of accepting new practices is medium, but the impact on the goal is significant). For goal \(G_{10}\) the migration from the old to a new customer support system (A13*) was a perceived as a serious threat to the goal’s success.

For identified risky goals a critical GQM*Strategies sub-grid can be plotted. Figure 24 illustrates the critical sub-grid of BigCorp’s grid. Goals \(G_6\) and \(G_{10}\) are risky goals (shaded goals in the figure). The critical sub-grid is comprised of the risky goals and goals that are connected with them (i.e. goals: \(G_4,G_2,G_1,G_0\)). The relationships among goals in a critical sub-grid can be interpreted as:
Table 28. Goals risk analysis.

<table>
<thead>
<tr>
<th>Goal</th>
<th>Key</th>
<th>Likelihood</th>
<th>Goal</th>
<th>Estimated</th>
<th>Risky</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[Type]</td>
<td>(Certainty)</td>
<td>Impact</td>
<td>RE</td>
<td>Cost</td>
</tr>
<tr>
<td>$G_0$</td>
<td>A10[*]</td>
<td>low</td>
<td>moder.</td>
<td>L</td>
<td>large</td>
</tr>
<tr>
<td>$G_1$</td>
<td>A10[*]</td>
<td>low</td>
<td>signif.</td>
<td>M</td>
<td>small</td>
</tr>
<tr>
<td>$G_2$</td>
<td>A11[*]</td>
<td>med.</td>
<td>moder.</td>
<td>M</td>
<td>med.</td>
</tr>
<tr>
<td>$G_3$</td>
<td>{C5,A4}[*]</td>
<td>low</td>
<td>signif.</td>
<td>M</td>
<td>small</td>
</tr>
<tr>
<td>$G_4$</td>
<td>A12[*]</td>
<td>low</td>
<td>moder.</td>
<td>L</td>
<td>small</td>
</tr>
<tr>
<td>$G_5$</td>
<td>{A9,A10}[+]</td>
<td>med.</td>
<td>low</td>
<td>L</td>
<td>small</td>
</tr>
<tr>
<td>$G_6$</td>
<td>A12[*]</td>
<td>med.</td>
<td>signif.</td>
<td>H</td>
<td>small</td>
</tr>
<tr>
<td>$G_7$</td>
<td>A12[*]</td>
<td>low</td>
<td>moder.</td>
<td>L</td>
<td>small</td>
</tr>
<tr>
<td>$G_8$</td>
<td>A12[*]</td>
<td>low</td>
<td>moder.</td>
<td>L</td>
<td>small</td>
</tr>
<tr>
<td>$G_9$</td>
<td>{A9,A10}[+]</td>
<td>low</td>
<td>moder.</td>
<td>L</td>
<td>small</td>
</tr>
<tr>
<td>$G_{10}$</td>
<td>A13[-]</td>
<td>high</td>
<td>signif.</td>
<td>VH</td>
<td>small</td>
</tr>
</tbody>
</table>

- The impact of risky goals on other goals. Risky goals exceed the level of acceptable risk, and failing to achieve those goals will have negative impacts on the connected upper-level goals. For example, failing to achieve $G_{10}$ can cause serious problems to achieving goal $G_1$, and if goal $G_1$ is not achieved, then the top-level goal $G_0$ will not be achieved. A goal failure does not imply total failure of the connected upper-level goals. In other words, the failure will cause deviations from the magnitude (goal target) of the connected upper-level goals. In the case of BigCorp, failing to increase sales 20% CAGR (goal $G_1$) in five years, can cause a smaller increase in net earnings ($G_0$) than originally planned.

- The importance of other goals in the critical sub-grid for the risky goals. The risky goals require more resources, i.e. they have high costs, and when compared with the benefits that they will provide in the given time period, the ratio cannot be justified with the goals risk level, meaning that the expected benefits are too low and the risk is too high. Therefore, the risky goals need upper-level goals to justify the investments required for their implementation. For example, goal $G_6$ is risky because the costs of establishing collective code ownership are more than three times the expected benefits. The collective code ownership is necessary to establish agile processes ($G_4$). However, the expected benefits of agile processes are high enough to justify the necessary investment in the $G_4$. Furthermore, if risky goal $G_4$ is achieved, but
upper-level goal $G_0$ fails, then the R&D unit might end up financing a successful goal $G_4$ that will never return the investment because the $G_0$ failed.

### 2.4.3 Define Cost–Benefit GQM Graph

In order to enable tracking of the costs and benefits, cost-benefit graph was defined as explained in Section 7.2.1.

### 2.4.4 Strategy Assessment and Grid Embedded Risk

For each simple i-strategy (Table 29), assessment card was created. The assessment card had two sides. On the left side, a single i-strategy was presented with the related goal and all refined goals with the strategy. In addition, all context and assumption elements that were relevant for presented strategy and goals were given. The right side of the card (Figure 36) was used to collect the feedback from an expert regarding the strategy effectiveness. The first question was about the likelihood of the effectiveness of the given strategy if all derived goals were already achieved. The second question was used to fine-tune the answer from the first question. The scales for both questions were designed to conveniently obtain probability distribution markers.

Participants were assessing only simple strategies that they developed. Finally, assessment for a single strategy represents a group assessment achieved after reaching consensus about strategy effectiveness. The same procedure was repeated for bottom-level goals (Table 29).

Using inputs provided by participants allows definition of probability distribution for BigCorp’s grid in the form of probability distribution tables, as depicted in Figure 44.

For identified risky goals the size of grid embedded risk is $\mathcal{R} = 0.00012$. That value is taken as a baseline for monitoring grid embedded risk during the grid execution.

Based on the estimates of costs and benefits for BigCorp’s goals and strategies, budget was defined. Figure 25 shows planned budget (i.e. estimated costs) and expected benefits realization plan.
Table 29. BigCorp example: results of the strategies assessment.

<table>
<thead>
<tr>
<th>Simple i-strategy</th>
<th>Success Likelihood</th>
<th>Assessment Fine Tuning</th>
<th>PD marker ( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_1 )</td>
<td>High</td>
<td>Optimistic</td>
<td>0.9</td>
</tr>
<tr>
<td>( \sigma_2 )</td>
<td>High</td>
<td>Neutral</td>
<td>0.8</td>
</tr>
<tr>
<td>( \sigma_3 )</td>
<td>Medium</td>
<td>Optimistic</td>
<td>0.6</td>
</tr>
<tr>
<td>( \sigma_4 )</td>
<td>Low</td>
<td>Neutral</td>
<td>0.2</td>
</tr>
<tr>
<td>( \sigma_5 )</td>
<td>Medium</td>
<td>Neutral</td>
<td>0.5</td>
</tr>
<tr>
<td>( \sigma_6 )</td>
<td>Medium</td>
<td>Optimistic</td>
<td>0.6</td>
</tr>
<tr>
<td>( \sigma_7 )</td>
<td>High</td>
<td>Optimistic</td>
<td>0.9</td>
</tr>
<tr>
<td>( \sigma_8 )</td>
<td>High</td>
<td>Pessimistic</td>
<td>0.7</td>
</tr>
<tr>
<td>( \sigma_9 )</td>
<td>High</td>
<td>Optimistic</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 30. BigCorp example: bottom-level goals (leaf nodes).

<table>
<thead>
<tr>
<th>Goal</th>
<th>Success Likelihood</th>
<th>Assessment Fine Tuning</th>
<th>PD marker ( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G_1 )</td>
<td>High</td>
<td>Neutral</td>
<td>0.8</td>
</tr>
<tr>
<td>( G_{10} )</td>
<td>Medium</td>
<td>Pessimistic</td>
<td>0.4</td>
</tr>
<tr>
<td>( G_6 )</td>
<td>Medium</td>
<td>Neutral</td>
<td>0.5</td>
</tr>
<tr>
<td>( G_7 )</td>
<td>High</td>
<td>Optimistic</td>
<td>0.9</td>
</tr>
<tr>
<td>( G_8 )</td>
<td>High</td>
<td>Pessimistic</td>
<td>0.7</td>
</tr>
<tr>
<td>( G_9 )</td>
<td>High</td>
<td>Optimistic</td>
<td>0.9</td>
</tr>
</tbody>
</table>

2.5 Execute Process

The BigCorp’s grid is executed, i.e. strategies are applied, data collected and validated, and feedback sessions performed. The execution of strategies is done by different organizational units, and each unit performs three activities during the grid execution:

- Process Execution. This represents implementation of strategies for local goals that are specified by the organizational unit;
- Analyze Results. The strategy implementation plans are monitored and analyzed. If needed, some local changes are made. In other words, changes that do not require reviewing upper-level goals and strategies are acceptable;
- Provide Process with Feedback. The knowledge and experience is packaged for sharing with other units and departments within the organization.

2.5.1 Process Execution

Plans for implementing unit specific strategies and goals are conducted, data collection occurs according to measurement plans. Issues that occurred are noted done for regular weekly follow up meetings, where they are handled and resolved. An important practice of issue handling it classification. It is important to identify if the organizational unit can handle an issue locally or it needs to be propagated further to higher levels in the organizational hierarchy.
2.5.2 Analyze Results

Earned Value Analysis

Budget and planned benefits do not just specify the total amount of the financial resources available and overall gains from benefits; they also specify the dynamics of expenditures and of benefits realization materialization.

For the BigCorp’s case, we created the budget and planned benefits according to the estimates assessed during the BVA. In order to demonstrate different potential situations that can occur during the execution phase, four different potential scenarios for MHP business unit are presented. The MHP sub-grid contains three goals: $G_2$, $G_5$, and $G_9$. Where for the MHP business unit the upper-level goal $G_2$ represents top-level goal, and for the purpose of discussing different scenarios it will be referred to as $B_1$. Other two goals $G_5$ and $G_9$ will be referred to as $B2$ and $B3$.

**Ideal case (Scenario 1)** In an ideal case, the goals realization proceeds according to plan with minor variations in the budgeted costs and planned benefits realization. Cost-related earned value metrics have similar values, i.e. cost and schedule variances are minimal. If visualized, such graph would depict plots that are close to each other.

The cost performance index for all goals converges toward a value of 1, meaning that all costs are on budget. It can be observed the same situation with benefits-related earned value metrics. Both performance indices are converging toward a value of 1, forming a kind of funnel shape.

**Total disaster (Scenario 2)** The second scenario represents the worst possible situation. Actual costs exceed budgeted costs, while there is no sign of goal realization. This situation is easily recognizable; $BCWP$ and $PBRM$ values are very small, close to zero, meaning that there is no progress with goals realization. Furthermore, because $BCWP < BCWS < ACWP$ is true for all goals and the gap between them is constantly increasing over time, actual cost is far over budget with uncorrectable delays. Cost performance indices converge toward small values, while the situation with the benefits performance indices is mirrored (larger values are an indication of benefits that are less materialized than planned).
Effective realization (Scenario 3)  This scenario illustrates a situation where the chosen strategies at all levels are achieving results, with limited but acceptable variations from the plan. Earned value metrics are plotted in Figure 45. The top-level business goal B1 was realized in the given timeframe of 5 years \((BCWS = BCWP\) in year 5), with around 20% more costs than budgeted (cost performance indices are converging toward 0.8). At the same time, there was about 50% less planned benefits realized (performance indices converge toward 1.5). If we further analyze the metrics, we can see that significant deviation from the plans occurred after the first year. Earned value metrics for the goal B3 indicate that realization of the goal was lagging behind in the first year \((BCWP < BCWS\)). Also, for the same period, the goal B3 was over budget \((ACWP > BCWP)\), and after one year and a half the goal B3 was achieved \((BCWS = BCWP)\). From the funnel-like shape formed by performance indices it can be conclude that in general all chosen strategies were achieving results. In the end, actual expenditures were about 592 FTEs, while realized benefits were about 795 FTEs. This situation illustrates the advantage of having information regarding the tracking of not only costs, but also of benefits. The possibility of balancing cost-related information with benefits realization increases the validity of the decision to proceed with goal realization despite over-budgeted costs.

Ineffective strategy (Scenario 4)  The last scenario illustrates a situation where realization of the lower-level goal proceeds according to plan, but at the same time, the upper-level goal realization is not achieved. Earned value metrics for this scenario are plotted in Figure 46. The plots for the top-level business goal B1 indicate that after the second year, the goal realization was halted (no changes in \(BCWP\) and \(PBRM\) values). With further analysis, we can conclude that the B3 goal was achieved after 1.5 years with additional costs of about 35% over budget \((CPI = 0.66)\). Further, on the next upper level, we can see that the realization of goal B2 was halted after the first year. The unsuccessful realization of B2 has consequences for the realization of the top-level goal B1. If we compare performance indices for B3 and B2, it is clear that they are diverging from each other. Apparently, the strategy that links B3 and B2 has to be reevaluated.

Risk Monitoring

Every six months, the key assumptions of the business goals were checked and the risk reassessed. For example, in scenario 3, the earlier identified threat of failing to
Fig 45. Scenario 3: earned value metrics for tracking budgeted cost and planned benefits realization of business goals $B_1$, $B_2$, and $B_3$. Costs-and-benefits performance indices.
Fig 46. Scenario 4: earned value metrics for tracking budgeted cost and planned benefits realization of business goals $B_1$, $B_2$, and $B_3$. Costs-and-benefits performance indices.
implement some key practices occurred. That was the reason for the delayed realization of the goals, as the product team could not work to full capacity. After solving that problem, the planned team of engineers was reinforced with additional members. This situation is visible from Figure 45 as the over-resourcing of goal B3 after the first year ($ACWP > BCWP$).

In addition, during the grid execution goal owners can feed in new estimates and predications regarding the effectiveness of strategies. These estimates allow for redefining the probability distribution and to recalculate new grid embedded risk. Figure 47 plots values of grid embedded risk for the period of five years. Initial estimate was taken as baseline.

Using earned value metrics, we can identify situations, like in our example, when goal realization is behind schedule, or when the goal is under- or over-resourced. Understanding of the reasons for such situations is provided by risk monitoring. In this example, the reason was not having specified product requirements on time.

2.5.3 Feedback

Interesting issues and problems resolved by a unit are shared and presented in monthly follow-up meetings. There are scheduled regular monthly meeting for presenting results of the grid implementation. The experiences are consolidated to formulate lessons learned.
2.6 Analyze Results

This step provides an opportunity to analyze data and revise strategies if needed. Findings are shared and communicated.

2.7 Package and Improve

This step is to adapt and improve the grid, correct wrong assumptions, and set new ones. Strategies should be adapted if assumption changes require that.
# Appendix 3  Strategy Assessment Follow-up Survey

## The Assessment Quality Questionnaire

*(to be completed after the strategy assessment exercise)*

### Q1: Providing information that is required for the assessment (i.e. answering questions Q1 and Q2 in the assessment cards), I find it (select the most appropriate answer):

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) [ ]</td>
<td>Extremely challenging—a significant mental effort is required and I lack confidence in the assessment results.</td>
<td></td>
</tr>
<tr>
<td>2) [ ]</td>
<td>Challenging but doable—a mental effort is required, but I am confident with the assessment results.</td>
<td></td>
</tr>
<tr>
<td>3) [ ]</td>
<td>Not too difficult and not too easy—e.g., there are other decisions in my work that require more challenging analysis or assessments than this one.</td>
<td></td>
</tr>
<tr>
<td>4) [ ]</td>
<td>Easy—It required an average mental effort and I am very comfortable with my assessment.</td>
<td></td>
</tr>
</tbody>
</table>

### Q2: My familiarity with the context of the given strategies and related goals is (select one possible answer):

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) [ ]</td>
<td>I am well versed with the context (environment) of the given strategies and related goals.</td>
<td></td>
</tr>
<tr>
<td>2) [ ]</td>
<td>I know enough about the context (environment) of the given strategies and related goals that I can provide a confident assessment.</td>
<td></td>
</tr>
<tr>
<td>3) [ ]</td>
<td>I have some understanding of the context (environment) of the given strategies and related goals, but not sufficient enough to provide a confident assessment.</td>
<td></td>
</tr>
<tr>
<td>4) [ ]</td>
<td>I am not familiar with the context (environment) of the given strategy and related goals.</td>
<td></td>
</tr>
</tbody>
</table>

### Q3: During the assessment my approach was (select the most appropriate answer):

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) [ ]</td>
<td>I analyzed the strategy and goals specifically for the given context and assumptions.</td>
<td></td>
</tr>
<tr>
<td>2) [ ]</td>
<td>I analyzed the strategy and goals specifically for the given context and assumptions, and I also identified other strategies for the same goal and did my evaluation relative to that strategy set.</td>
<td></td>
</tr>
<tr>
<td>3) [ ]</td>
<td>I relied more on my feeling rather than on some rational reasoning behind my assessment.</td>
<td></td>
</tr>
<tr>
<td>4) [ ]</td>
<td>None of those.</td>
<td></td>
</tr>
</tbody>
</table>

![Fig 48. The strategy assessment follow-up survey.](image)
Appendix 4 Interview Guide: Assessment of the Strategy Effectiveness

4.1 Interview Questions

Q1: **How difficult was it for you to give the assessment?** [The focus is on the assessment task and how the difficulty of the task impacted the validity of results, i.e. strategy assessment]

✓ Can you identify why it was difficult?
✓ How much mental effort you had to put in the assessment?
✓ Are you confident with your assessment? If not, why not?

Q2: **How familiar are you with the context of the given strategy and related goals?** [The focus is on the interviewee’s familiarity with the context, and how that impacts the assessment.]

✓ Are you familiar enough with the context to give confident assessment?
✓ In your opinion, how familiar should someone be with the context in order to provide a good assessment? [Nobody can have absolute knowledge about the context and environment; it will always require some assumptions and predictions]

Q3: **Can you describe what your approach (i.e. mental process) was during the assessment?** [The question tries to elicit a mental model that interviewee used to provide the assessment.]

✓ Have you tried to analyze uncertainty in connection with the given strategy?
✓ Have you tried to identify alternative strategies for the given strategy?
✓ Have you relied more on your feelings than rational thinking?

Q4: **Does this kind of strategy evaluation (i.e. assessment of the strategy effectiveness) help you to better understand the strategy and identify other relevant strategies that should be considered?** [The focus of the question is on the assessment process and questioning if the process itself, despite the assessment results, has a value (i.e. is useful).]

✓ Should all strategies in a grid be evaluated in this way during grid development?
Could recording the evaluation process help others understand the grid better?

If you did this as a group activity would it help improve the understanding of all the people involved? Is this process useful for identifying other strategies that might be better than the existing ones in the grid?
Appendix 5 Summary of the Interview Results of the Strategy Assessment Study.

Table 31. Summary of the interview notes. Expert type can be I–industry or A–academia.

<table>
<thead>
<tr>
<th>Expert Type</th>
<th># of cards</th>
<th>Q1. Difficulty</th>
<th>Q2. Familiarity with context</th>
<th>Q3. Mental process</th>
<th>Q4. Usefulness and suggestions</th>
</tr>
</thead>
<tbody>
<tr>
<td>001 A</td>
<td>2</td>
<td>Challenging but doable. Once when the situations are abstracted with assumptions the assessment is feasible.</td>
<td>Familiar enough to be able to generalize and abstract the environment.</td>
<td>Focusing on given strategy and imagining potential scenarios for that strategy. The assessment is based on those scenarios.</td>
<td>It is useful to indicate the level of agreement toward specific strategies within a group of people.</td>
</tr>
<tr>
<td>002 A</td>
<td>2</td>
<td>Context and assumptions were not sufficient for giving the assessment, because it was an artificial example.</td>
<td>(Observation made by expert) There are two scenarios: (1) the grid should contain all information needed to make an assessment or (2) the grid should contain the most important information, so an expert should rely on his knowledge of the context in order to make an assessment.</td>
<td>Just analyzed information presented in the assessment card.</td>
<td>It can be used as a tool for checking the belief of the wider group of people in an organization, after a grid has been defined.</td>
</tr>
<tr>
<td>Expert Type ID.</td>
<td># of cards</td>
<td>Q1. Difficulty</td>
<td>Q2. Familiarity with context</td>
<td>Q3. Mental process</td>
<td>Q4. Usefulness and suggestions</td>
</tr>
<tr>
<td>----------------</td>
<td>------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>003</td>
<td>4</td>
<td>Challenging but doable. Need to consider the external factors that might change. Some strategies were from higher levels and required higher level thinking.</td>
<td>At lower level well versed. For the higher level goals and strategies some understanding of the context.</td>
<td>Tried to see company needs as a whole. Tried to make the multi-level connections of the strategy. Did not look for alternate strategies but study what was there.</td>
<td>The assessment is important and needed while changing the grid, to identify which strategies need to be modified. Good to do as a group activity, so you can hear and include the opinions of the others. You can do better assessment after some time when you get a better feel for what is happen.</td>
</tr>
<tr>
<td>004</td>
<td>4</td>
<td>It was very easy. Well versed with the context of the organizational goals and strategies.</td>
<td></td>
<td>The assessment was built upon the knowledge regarding the goals and strategies. (The expert was intensively involved in building the grid.)</td>
<td>No suggestions.</td>
</tr>
<tr>
<td>Expert Type ID</td>
<td># of cards</td>
<td>Q1. Difficulty</td>
<td>Q2. Familiarity with context</td>
<td>Q3. Mental process</td>
<td>Q4. Usefulness and suggestions</td>
</tr>
<tr>
<td>----------------</td>
<td>------------</td>
<td>----------------</td>
<td>----------------------------</td>
<td>-------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>005 I 3</td>
<td></td>
<td>The assessment was easy. The expert was confident with the assessment not only because of familiarity with the context and environment, but also because of time distance. Some goals and strategies were defined over six months before the assessment, and passed time confirmed some assumptions that were stated at the beginning.</td>
<td>Familiarity with context and environment helps to make better assessments. The grid should document just enough information that someone who is familiar with the environment can provide the assessment.</td>
<td>Based on analyzes of the information presented in the assessment card, the presented model (goals, strategies, context and assumptions) was compared by the model in expert's head. Based on that comparison strategies were assessed.</td>
<td>Good tool for checking &quot;quality&quot; of strategies and goals. Can indicate missing of insufficient information. It can generate good discussion points if the strategies are reviewed by a group of people. It can indicate a level of agreement toward specific strategies within organization.</td>
</tr>
<tr>
<td>006 A 2</td>
<td></td>
<td>Not too difficult and not too easy. Familiar enough with the context.</td>
<td></td>
<td>Analyzed the strategies and goals for the given context and assumptions.</td>
<td>No suggestions.</td>
</tr>
</tbody>
</table>
582. Haataja, Tatu (2011) Peroxisomal multifunctional enzyme type 2 (MFE-2) : the catalytic domains work as independent units
584. Hautala, Johanna (2011) Academic knowledge creation as a spatio-temporal process : The case of international research groups in Finland
587. Rajanen, Mikko (2011) Applying usability cost-benefit analysis — explorations in commercial and open source software development contexts
588. Tervo, Heli (2011) Information technology incidents in the present information society : Viewpoints of service providers, users, and the mass media
589. Riipinen, Katja-Anneli (2011) Genetic variation and evolution among industrially important Lactobacillus bacteriophages
590. Lampila, Petri (2011) Populations and communities in human modified forest landscapes
591. Liukkunen, Kari (2011) Change process towards ICT supported teaching and learning
593. Tiikkaja, Marjo (2012) Value creation in collaboration between software suppliers and customers: suppliers’ perspective
596. Härkönen, Laura (2012) Seasonal variation in the life histories of a viviparous ectoparasite, the deer ked

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/
Vladimir Mandić

MEASUREMENT-BASED VALUE ALIGNMENT AND REASONING ABOUT ORGANIZATIONAL GOALS AND STRATEGIES

STUDIES WITH THE ICT INDUSTRY